



# Analyse en Ondelettes et par Paquets d'Ondelettes de Processus Aléatoires Stationnaires, et Application à l'Estimation Non-Paramétrique

Abdourrahmane Atto

## ► To cite this version:

Abdourrahmane Atto. Analyse en Ondelettes et par Paquets d'Ondelettes de Processus Aléatoires Stationnaires, et Application à l'Estimation Non-Paramétrique. Mathématiques [math]. Ecole Nationale Supérieure des Télécommunications de Bretagne - ENSTB; Université Rennes 1, 2008. Français. NNT: . tel-00384852

**HAL Id: tel-00384852**

**<https://theses.hal.science/tel-00384852>**

Submitted on 16 May 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

présentée à

**TELECOM BRETAGNE**

Sous le sceau de l'**Université Européenne de Bretagne**

pour obtenir

le grade de : **DOCTEUR DE TELECOM BRETAGNE,**

**EN CO-DÉLIVRANCE AVEC L'UNIVERSITÉ DE RENNES 1**

Mention : *Mathématiques et Applications*

par

**Abdourrahmane Mahamane Atto**

Équipe d'accueil : Département Signal & Communication, Institut TELECOM - TELECOM Bretagne

---

## Analyse en Ondelettes et par Paquets d'Ondelettes de Processus Aléatoires Stationnaires, et Application à l'Estimation Non-Paramétrique

---

Soutenue le 30 Septembre 2008 devant la commission d'Examen :

### Composition du Jury :

- Présidente :* Karine TRIBOULEY, Professeur, Université Paris X - Nanterre
- Rapporteurs :* Anestis ANTONIADIS, Professeur, Université Joseph Fourier - Grenoble  
Laure BLANC-FERAUD, Directeur de recherche, INRIA - Sophia-Antipolis
- Examineurs :* Rémi GRIBONVAL, Chargé de recherche, IRISA - Rennes  
Alain HILLION, Professeur, TELECOM Bretagne - Brest  
Dominique PASTOR, Maître de conférences, TELECOM Bretagne - Brest
- Invité :* Grégoire MERCIER, Maître de conférences, TELECOM Bretagne - Brest



La solution à un problème très complexe peut se jouer sur quelques détails<sup>1</sup>.

---

1. Enseignement relatif à la persévérance et à l'humilité, tiré du proverbe hausa "*le mensonge peut courir un an, la vérité le rattrape en un jour*".



---

# Remerciements

Le travail présenté dans ce mémoire de thèse est le résultat de trois années d'études sur un thème de recherche qui nécessite des développements théoriques et dont l'objectif est de permettre l'amélioration des performances en estimation non-paramétrique de signaux et d'images. Ce travail a été réalisé grâce à de nombreuses contributions.

Je tiens d'abord à adresser mes remerciements à Dominique Pastor pour la supervision de ce travail. Par son dynamisme et sa recherche permanente de solutions, il a réussi à orienter ce travail vers une analyse cohérente de certains problèmes fondamentaux liés aux décompositions par paquets d'ondelettes de processus aléatoires, sans perdre de vue l'objectif final qui est l'intérêt pratique de cette analyse.

Mes remerciements s'adressent aussi à Grégoire Mercier pour sa participation à l'encadrement ce travail. La perspicacité de son point de vue m'a conduit à la synthèse de la famille de fonctions à atténuations sigmoïdales.

Je remercie également Alexandru Isar pour ses conseils, ainsi que sa participation active dans les discussions autour de certaines questions fondamentales traitées dans cette thèse.

Mes remerciements vont également à l'endroit de Alain Hillion qui a accepté de diriger ce travail malgré toutes ses charges (énormes) de Directeur scientifique de TELECOM Bretagne et Directeur du Lab-STICC (Laboratoire en Sciences et Technologies de l'Information, de la Communication et de la Connaissance).

D'autre part, mes remerciements vont aussi à l'endroit des membres du département "Signal et Communication" pour les diverses contributions qu'ils m'ont apportées, ainsi que pour l'ambiance chaleureuse qui caractérise cette équipe.

Enfin, je veux exprimer mes remerciements à ma famille (au sens large) pour - entre autres - le soutien qu'ils m'ont apporté et la patience dont ils ont fait preuve durant ces trois longues années de recherche.



---

# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>I Propriétés statistiques</b>	<b>5</b>
<b>Introduction</b>	<b>7</b>
<b>1 Généralités</b>	<b>11</b>
1.1 Ondelettes et bancs de filtres . . . . .	11
1.2 $M$ -TPOD : Principe de la décomposition . . . . .	12
1.3 Représentation des chemins de l'arbre de la $M$ -TPOD . . . . .	14
1.4 $M$ -TPOD de Shannon associée à l'espace de Paley-Wiener . . . . .	16
<b>2 Analyse asymptotique de la séquence des fonctions d'autocorrélations</b>	<b>19</b>
2.1 $M$ -TPOD des processus aléatoires stationnaires . . . . .	19
2.2 Autocorrélations asymptotiques de la $M$ -TPOD de Shannon . . . . .	21
2.2.1 Convergence des fonctions d'autocorrélations . . . . .	21
2.2.2 Vitesse de convergence du processus de décorrélation . . . . .	23
2.3 Analyse asymptotique pour quelques familles de filtres . . . . .	24
2.3.1 Convergence des fonctions d'autocorrélations . . . . .	24
2.3.2 Résultats expérimentaux . . . . .	26
2.3.3 Convergence vers la variance du processus limite . . . . .	28
<b>3 Analyse asymptotique des distributions</b>	<b>33</b>
3.1 Rappels - État de l'art . . . . .	33
3.2 Distributions asymptotiques des coefficients de la $M$ -TPOD . . . . .	35
3.3 Résultats expérimentaux . . . . .	37



<b>Conclusion</b>	<b>41</b>
<b>II Estimation non-paramétrique</b>	<b>43</b>
<b>Introduction</b>	<b>45</b>
<b>4 Estimation non-paramétrique - Rappels et état de l'art</b>	<b>49</b>
4.1 Parcimonie de la représentation par ondelettes . . . . .	50
4.2 Estimation non-paramétrique par seuillage doux . . . . .	50
<b>5 Les seuils de détection pour l'estimation non-paramétrique</b>	<b>57</b>
5.1 Les seuils de détection . . . . .	57
5.2 Performances des seuils de détection . . . . .	62
5.2.1 Résultats expérimentaux sur les signaux . . . . .	62
5.2.2 Résultats expérimentaux sur les images . . . . .	64
5.3 Débruitage par transformée redondante . . . . .	66
5.3.1 Débruitage de signal . . . . .	66
5.3.2 Débruitage d'images . . . . .	66
5.3.3 Amélioration des performances . . . . .	69
<b>6 Estimation non-paramétrique par atténuation sigmoïdale</b>	<b>71</b>
6.1 Contexte de la méthode . . . . .	71
6.2 Définition, propriétés des fonctions SSBS . . . . .	73
6.3 Régularisation avec les fonctions SSBS . . . . .	76
6.4 Sélection des paramètres des fonctions SSBS . . . . .	77
6.4.1 Optimisation SURE des paramètres des fonctions SSBS . . . . .	77
6.4.2 Choix heuristique des paramètres des fonctions SSBS . . . . .	80
6.5 Débruitage SSBS du bruit coloré . . . . .	85
<b>Conclusion</b>	<b>89</b>
<b>Conclusion Générale et Perspectives</b>	<b>91</b>
<b>A Représentations dans les espaces de Hilbert</b>	<b>103</b>
A.1 Rappels d'analyse . . . . .	103
A.2 Représentations dans les Espaces de Hilbert . . . . .	107
A.2.1 Caractérisation d'une fonction par ses coefficients de frames . . . . .	108
A.2.2 Frame ajustée - Base de Riesz . . . . .	108

---

A.2.3 Représentations en ondelettes . . . . .	109
<b>B Bornes de l'opérateur de frames</b>	<b>111</b>



---

# Abbreviations

---

---

<b>M-TPOD</b>	Transformation par Paquets d'Ondelettes Discrètes à $M$ -bandes
<b>TPOD</b>	Transformation par Paquets d'Ondelettes Discrètes (2-bandes)
<b>TOD</b>	Transformation en Ondelettes Discrètes
<b>AR</b>	Auto-Régressif (processus)
<b>GG</b>	Gaussien Généralisé (processus, distribution)
<b>KS</b>	Kolmogorov-Smirnov (test statistique)
<b>EQM</b>	Erreur Quadratique Moyenne
<b>RSB</b>	Rapport Signal à Bruit

---

---

<b>SSBS</b>	Smooth Sigmoid-Based Shrinkage (famille de fonctions)
<b>NNG</b>	Non-Negative Garrote (fonction)
<b>SCAD</b>	Smoothly Clipped Absolute Deviation (fonction)
<b>SWT</b>	Stationary Wavelet Transform (version redondante de la TOD)
<b>SWPT</b>	Stationary Wavelet Packet Transform (version redondante de la TPOD)
<b>PSNR</b>	Peak Signal-to-Noise Ratio
<b>LD</b>	Level-Dependent (stratégie)
<b>SURE</b>	Stein's Unbiased Risk Estimate
<b>SURELET</b>	SURE - Linear Expansion of Thresholds (méthode)
<b>BLS-GSM</b>	Bayes Least Squares - Gaussian Scale Mixtures (méthode)

---

---



---

## Liste des figures

1.1	Arbre de la décomposition en 2-TPOD de $\mathbf{U}$ ( $j = 3$ ). . . . .	14
1.2	Arbre de la décomposition en 3-TPOD de $\mathbf{U}$ ( $j = 3$ ). . . . .	14
1.3	Ordre fréquentiel des paquets de l'arbre de la 2-TPOD de Shannon. . . . .	17
1.4	Ordre fréquentiel des paquets de l'arbre de la 3-TPOD de Shannon. . . . .	18
2.1	Réponses fréquentielles des filtres d'échelles de Daubechies. . . . .	26
2.2	Spectres et fonctions d'autocorrélation normalisées. . . . .	27
2.3	Fonctions d'autocorrélation normalisées ( $j = 3, 6$ , $r = 1, 7$ et $\mu = 0.5$ ). . . . .	30
2.4	Fonctions d'autocorrélation normalisées ( $j = 3, 6$ , $r = 1, 7$ et $\mu = 0.9$ ). . . . .	31
2.5	Convergence de la variance des coefficients vers la variance limite. . . . .	32
3.1	Distributions GGs de variance unité et de formes $\beta = 2, 1, 0.5$ . . . . .	38
3.2	Histogrammes des coefficients de la TPOD ( $j = 6, r = 7$ ). . . . .	41
4.1	Exemples de signaux tests, avec leurs représentations par TOD. . . . .	51
4.2	Quelques images naturelles, ainsi que leurs représentations par TOD. . . . .	52
4.3	Représentations par TOD de l'image 'Lena' et de ses versions bruitées. . . . .	53
4.4	Graphes des fonctions de seuillage dur, doux, NNG et SCAD. . . . .	54
5.1	Débruitage du signal 'Doppler' par SWT 'sym8' avec $J = 6$ et par seuillage doux. . .	67
5.2	Zoom du débruitage du signal 'Doppler'. . . . .	68
5.3	Image bruitée et images débruités de 'Lena'. . . . .	69
6.1	Histogrammes des coefficients de détails verticaux. . . . .	72
6.2	Histogrammes obtenus après seuillage. . . . .	73
6.3	Graphes des fonctions $\delta_{t,\tau,\lambda}$ pour $t = 0$ et $t \neq 0$ . . . . .	74
6.4	Graphe de la fonction $\delta_{\tau,\lambda}$ dans le demi-plan positif. . . . .	75
6.5	Allures des fonctions SSBS et de leurs fonctions de pénalité. . . . .	77
6.6	Image 'Lena' corrompue par du bruit additif blanc Gaussien d'écart-type $\sigma = 35$ . .	80

6.7 Débruitage de l'image 'Lena' corrompue par du bruit additif blanc Gaussien. . . .	81
6.8 Histogrammes des coefficients obtenus après atténuation SSBS. . . . .	85
6.9 Débruitage du signal 'Doppler' corrompu par du bruit rose. . . . .	87
6.10 Débruitage de l'image 'Lena' corrompue par du bruit de type speckle. . . . .	88

---

## Liste des tableaux

2.1 Valeurs de $\gamma(a(\kappa))$ pour les quatre séquences tests. . . . .	27
3.1 Convergence vers la distribution limite (processus GG). . . . .	39
3.2 Convergence vers la distribution limite (processus AR(1)-GG). . . . .	39
3.3 Convergence vers la distribution limite (processus AR(1)-GG). . . . .	40
5.1 Bornes supérieures $V(\sqrt{2\ln N})$ de la probabilité d'erreur du test à seuil $\mathcal{T}_{\lambda_D(N)}$ . . . . .	59
5.2 Seuils de détection, minimax, et universel pour différentes valeurs de $N$ . . . . .	60
5.3 Risques $r_\lambda$ pour les seuils de détection, minimax, et universel (signaux). . . . .	63
5.4 Valeurs $r_0(\mathbf{d})$ pour les images testées. . . . .	65
5.5 Risques $r_\lambda$ pour les seuils de détection, minimax, et universel (images). . . . .	65
5.6 Risques pour les seuils de détection, minimax, et universel (image 'Lena'). . . . .	67
5.7 PSNRs pour les seuils de détection, minimax, et universel. . . . .	69
6.1 Paramètres SSBS optimaux au sens SURE pour la TOD. . . . .	78
6.2 Paramètres SSBS et LD-SSBS optimaux au sens SURE pour la TOD. . . . .	79
6.3 PSNRs calculés avec les paramètres SURE pour le débruitage de l'image 'Lena'. . . . .	80
6.4 PSNRs moyens sur 10 réalisations du bruit additif blanc Gaussien. . . . .	83
6.5 PSNRs SWT et SSBS $_{\pi/K_0, \lambda_D(a_j, p_j^*)}$ calculés pour le débruitage de l'image 'Lena'. . . . .	84





---

# Introduction générale

## Motivations et contexte de l'étude

En traitement de signal et de l'information, les données observées sont très souvent composées de l'information utile (signal) corrompue par une perturbation généralement indépendante. La perturbation (bruit) peut être interne au système physique de traitement de l'information ou externe, due aux interactions entre le signal et les éléments de son milieu de propagation. La distribution du signal n'est pas toujours connue et elle s'avère assez souvent difficile à estimer en pratique. Il en est de même concernant la nature de la perturbation : sa distribution n'est pas connue et est assez souvent difficile à modéliser. Face à cette situation, la théorie statistique usuelle de la décision et de l'estimation basée sur l'emploi du rapport de vraisemblance ne s'applique pas rigoureusement. Il est donc nécessaire de développer des méthodes à coût minimum en termes de connaissances disponibles sur la nature des composantes de l'information reçue. On parle alors de décision et d'estimation (plus généralement de statistique) non-paramétriques au sens où l'on ne tient pas compte de connaissances *a priori* sur les distributions des signaux constituant l'observation. Une des méthodes les plus remarquables en statistique non-paramétrique trouve son fondement dans les propriétés de certaines bases fonctionnelles à pouvoir discriminer les différentes composantes d'une observation spécifique constituée de signaux de natures différentes. C'est ainsi que de nombreux travaux (voir [Donoho and Johnstone, 1994], entre autres) ont permis de mettre en évidence le fait suivant. La transformée en ondelettes possède une propriété dite de parcimonie de la représentation : l'énergie de signaux réguliers ou réguliers par morceaux se répartit sur un petit nombre de coefficients de grandes amplitudes. À l'opposé, la structure des composantes d'une certaine classe de perturbations reste sensiblement inchangée par cette même transformée. Cette capacité de compacter l'énergie de signaux tandis que la contribution du bruit demeure inchangée rend possible et performante l'estimation non-paramétrique par simple seuillage des coefficients du signal observé dans une base d'ondelettes. En plus de cette propriété de représentation parcimonieuse, les bases d'ondelettes possèdent une propriété de "décorrélation" qui a été mise en évidence dans plusieurs travaux (on en propose un résumé dans l'introduction de [Atto, Pastor, and Isar, 2007]). Cette propriété est de nature à favoriser l'utilisation des bases d'ondelettes pour simplifier divers problèmes de traitement de signaux, d'analyse des données et de théorie de l'information, compte tenu du fait que l'analyse de données non corrélées est souvent beaucoup plus simple que celle de données corrélées.

## Objectif et contributions

Le but visé à travers cette thèse est d'étudier plus en détail, les propriétés statistiques des transformations en ondelettes et par paquets d'ondelettes en vue de rendre encore plus simples et/ou plus performants, divers traitements de signaux. Plus particulièrement, l'application étudiée dans ce travail est l'estimation non-paramétrique de signaux observés en présence de bruit. Le programme est le suivant : d'une part, il s'agira d'analyser les propriétés qui permettront d'améliorer l'estimation non-paramétrique lorsque le bruit considéré est supposé être blanc et Gaussien (modèle de base) ; ces propriétés sont relatives à la nature parcimonieuse des représentations en ondelettes. D'autre part, il faudra étudier les conditions qui peuvent permettre d'améliorer le débruitage de bruit colorée ; conditions concernant la nature (décorrélation, loi) des coefficients d'ondelettes ou de paquets d'ondelettes de processus aléatoires et le type de traitement à effectuer pour supprimer ce bruit. Cette thèse apporte ainsi une contribution à l'étude des propriétés statistiques des transformations par paquets d'ondelettes [Atto, Pastor, and Isar, 2007], [Atto and Pastor, 2008b], [Atto and Pastor, 2008a], [Atto, Pastor, and Isar, 2006], et une contribution à l'amélioration des performances en estimation non-paramétrique de signaux observés en présence de bruit [Atto, Pastor, and Mercier, 2008a], [Atto, Pastor, and Mercier, 2008b], [Atto, Pastor, and Mercier, 2008c], [Pastor and Atto, 2008].

## Organisation du document

Le document est divisé en deux parties. Chaque partie est composée d'une introduction, de quelques chapitres spécifiques et d'une conclusion. Une conclusion générale, ainsi que les perspectives à ce travail sont données à la fin du document.

La première partie traite des propriétés statistiques des transformations en ondelettes et par paquets d'ondelettes. Elle présente les outils nécessaires à la compréhension de ces transformées, et les contributions apportées par cette thèse à l'étude de leurs propriétés statistiques. Les résultats énoncés sont basés sur un formalisme décrivant les chemins de l'arbre des paquets d'ondelettes sous forme de séquences à valeurs dans un alphabet à  $M$  éléments,  $M$  désignant le nombre de filtres à appliquer pour une étape de la décomposition. Ces résultats montrent que les coefficients associés à un chemin donné dans l'arbre des paquets d'ondelettes convergent vers un processus blanc et Gaussien dont la variance dépend de la succession de filtrage effectuée lorsqu'on parcourt ce chemin, et cela lorsque le niveau de résolution et la régularité des filtres d'analyse sont suffisamment grands.

La seconde partie traite de l'estimation non-paramétrique des signaux par analyse en ondelettes. Après un bref rappel sur le principe de cette estimation et une analyse de la parcimonie de la transformée en ondelettes, nous proposons l'utilisation de seuils dits de détection pour la distinction des coefficients significatifs du signal. Une analyse théorique et expérimentale des performances de l'estimation est alors effectuée dans le cadre de l'estimation des coefficients du signal par la fonction dite de seuillage doux, calibrée avec les seuils de détection. Cette analyse des performances des seuils de détection est réalisée en comparaison avec les performances des seuils classiques de la littérature, à savoir les seuils universels et minimax. Enfin, pour terminer cette partie, nous proposons une nouvelle famille de fonctions pour l'estimation

des coefficients du signal. Ces fonctions opèrent, de manière très régulière, une atténuation sigmoïdale sur les coefficients d'ondelettes. Elles dépendent de trois paramètres qui permettent de régler le degré d'atténuation imposée aux coefficients, afin de mieux tenir compte du degré de parcimonie de la représentation du signal étudié.



## **Première partie**

# **Propriétés statistiques des décompositions en ondelettes et par paquets d'ondelettes des processus aléatoires stationnaires**



---

# Introduction

Dans cette partie,  $M$  désigne un entier naturel,  $M \geq 2$ . Les atomes (fonctions de bases) d'une *Transformation par Paquets d'Ondelettes Discrètes* à  $M$ -bandes ( $M$ -TPOD) sont notées  $W_{j,n,k}$ , où  $j \in \mathbb{N}$  désigne le niveau de décomposition;  $n \in \{0, 1, 2, \dots, M^j - 1\}$  est le paramètre de décalage;  $k \in \mathbb{Z}$  est le paramètre de translation. La transformation par paquets d'ondelettes discrètes classique correspond au cas  $M = 2$  et est notée simplement TPOD au lieu de 2-TPOD. La *Transformation en Ondelettes Discrètes* (TOD) classique est réalisée à partir des fonctions  $W_{j,n,k}$  pour  $n \in \{0, 1\}$ .

Les TOD et  $M$ -TPOD s'implémentent par le biais de bancs de filtres à reconstruction parfaite, dits *filtres d'analyse* (encore appelé *filtres de décomposition*). Par la suite, les termes transformations ou représentations en ondelettes seront associés aux termes TOD, TPOD,  $M$ -TPOD, ainsi qu'à toutes leurs variantes redondantes. La précision spécifique à l'une de ces transformées sera apportée lorsque cela est nécessaire.

Les transformées en ondelettes se sont largement imposées dans l'analyse et la résolution de problèmes liés aux sciences et aux techniques de l'ingénierie. Cet essor est dû principalement à deux propriétés spécifiques qui résultent des décompositions de signaux et de processus aléatoires par projection sur les bases d'ondelettes : la *parcimonie* de représentation et le *blanchiment* de données corrélées.

La parcimonie d'une représentation de signal décrit la capacité de cette représentation à répartir l'énergie du signal sur un petit nombre de coefficients de grandes amplitudes. Les transformées en ondelettes permettent d'obtenir une représentation parcimonieuse de signaux réguliers et réguliers par morceaux en ce sens. La parcimonie des représentations TOD pour les signaux réguliers et réguliers par morceaux a été mise en évidence par les travaux de Donoho et Johnstone ([Donoho and Johnstone, 1994], [Johnstone, 1999], entre autres). La parcimonie d'une TPOD, pour ces mêmes classes de signaux, découle de celle de la TOD car : d'une part, la plupart des coefficients de grandes amplitudes sont les coefficients obtenus par projection sur  $(W_{j,0,k})_{k \in \mathbb{Z}}$  (coefficients d'approximation), les coefficients issus de la projection sur  $(W_{j,1,k})_{k \in \mathbb{Z}}$  (coefficients de détails) étant d'amplitudes beaucoup plus faibles. D'autre part, on obtient une TPOD à partir d'une TOD par décomposition récursive des  $(W_{j,1,k})_{k \in \mathbb{Z}}$  (décomposition des coefficients de détails), ce qui ne change pas significativement la nature parcimonieuse de la représentation obtenue.

En plus de la propriété d'admettre une représentation parcimonieuse pour les signaux réguliers et réguliers par morceaux, de nombreux travaux ont permis de mettre en évidence l'effet de blanchiment (propriété de *décorrélacion*) de la représentation de certains proces-



sus aléatoires par décomposition en ondelettes. Ces travaux ([Tewfik and Kim, 1992; Flandrin, 1992; Cohen, 1992; Dijkerman and Mazumdar, 1994; Walter, 1994; Houdré, 1994], entre autres) montrent que les coefficients inter-échelles et intra-échelles des TOD de certains processus stochastiques tendent à être décorrélés lorsque le niveau de décomposition augmente. La TOD se présente donc comme une alternative à la transformation de Karhunen-Loève, alternative commode en ce sens qu'elle ne nécessite pas de résolution d'équation aux valeurs propres et n'impose pas que le processus décomposé soit limité en temps, ce qui est le cas avec la transformée de Karhunen-Loève. Ainsi, [Lemoine and Pastor, 1998] propose un récepteur de télécommunication robuste en présence de bruit stationnaire et basé sur la TOD. Ce récepteur fonctionne sans connaissance *a priori* de la densité spectrale du bruit et permet d'obtenir des taux d'erreurs binaires comparables à ceux obtenus avec le récepteur optimal basé sur la transformée de Karhunen-Loève. D'autre part, on peut noter que lorsque la fonction d'auto-corrélation d'un processus aléatoire stationnaire au sens large est connue, il est possible de construire une base (non orthogonale) d'ondelettes sur laquelle le processus considéré peut être représenté avec une séquence de coefficients décorrélés. La construction se fait sans résolution d'équation aux valeurs propres et demeure valide que le processus soit en temps limité ou non [Zhang and Walter, 1994].

Notons cependant que la décorrélation due à la TOD est asymptotique avec le niveau de décomposition. Or, une séquence de coefficients de la TOD est constituée par la concaténation des séquences de coefficients d'ondelettes aux différents niveaux de décomposition considérés. Cette concaténation inclut les vecteurs des coefficients de détails situés aux premiers niveaux de décomposition. Ces derniers peuvent présenter une forte corrélation lorsque le processus décomposé est lui-même fortement corrélé. Ainsi, la TOD présente une limitation qui est liée au fait que les séquences des coefficients de détails situés aux premiers niveaux de décomposition, lorsqu'ils présentent une corrélation, vont demeurer irrémédiablement corrélés. Une solution consiste donc à re-décomposer ces mêmes séquences de coefficients de détails, ce qui revient au principe de base de la TPOD.

Plus généralement, une TPOD offre le choix de plusieurs bases, et permet entre autres de sélectionner des bases (donc des séquences de coefficients) toutes situées au-delà d'un certain niveau de décomposition. Ainsi, si la tendance au blanchiment des coefficients restait valide dans le cas de la TPOD, on pourrait alors décomposer un processus en séquences de coefficients, toutes assez sensiblement décorrélées.

On retrouve dans [Pastor and Gay, 1995], [Isar and Naforniță, 1998] et [Leporini and Pesquet, 1999], des résultats concernant le comportement asymptotique des fonctions d'autocorrélation issues de la décomposition par la TPOD de processus aléatoires stationnaires au sens large. Dans [Pastor and Gay, 1995], la décorrélation asymptotique des coefficients de la TPOD de processus aléatoires stationnaires au sens large est obtenue par le biais de filtres d'analyse de régularité suffisamment élevée. Dans [Isar and Naforniță, 1998], on retrouve une tentative de généralisation du processus de décorrélation de la TPOD par une démarche intuitive et analogue à celle utilisée pour démontrer la décorrélation asymptotique dans le cas de la TOD. Cette généralisation semble ne pas dépendre de la régularité des filtres d'analyse. La démarche utilisée dans [Isar and Naforniță, 1998] est justifiée (et est étendue au cas de  $M$ -TPOD) par [Leporini and Pesquet, 1999]. En fait, dans [Leporini and Pesquet, 1999], les atomes  $W_{j,n,k}$  de la décomposition en  $M$ -TPOD sont considérés avec un paramètre de décalage  $n$  supposé constant, et les résultats énoncés sont exactement les mêmes que

ceux obtenus dans le cas de la TOD : la décorrélation est asymptotique avec le niveau de décomposition, et les fonctions d'autocorrélation asymptotiques ne dépendent que de la valeur de la densité spectrale du processus aléatoire décomposé en 0, processus supposé du second ordre, stationnaire au sens large, et admettant une densité spectrale de puissance (spectre). Les résultats présentés dans [Isar and Naforniță, 1998], [Leporini and Pesquet, 1999] sont donc indépendants à la fois de la nature des filtres de décomposition utilisés et du chemin choisi dans l'arbre de la TPOD, et de ce fait, indépendants de la nature et de la combinaison des filtrages effectués (succession de filtrages passe-bas, passe-bande, passe-haut).

Or, assez récemment, [Craigmile and Percival, 2005] a montré le rôle joué par la longueur des réponses impulsionnelles des filtres de Daubechies pour obtenir des coefficients de TOD décorrelés (décorrélation inter-échelles) : la covariance de certains processus stationnaires décroît vers 0 de l'ordre de  $L^{-1/4}$  lorsque la longueur  $L$  des filtres de Daubechies augmente. La nature des filtres semble donc être importante dans le processus de décorrélation de la TOD. La TPOD étant encore plus complexe, l'analyse du processus de décorrélation de la TPOD est sans doute encore plus compliquée que la généralisation présentée dans [Isar and Naforniță, 1998] ou [Leporini and Pesquet, 1999].

Au cours de cette thèse, nous avons donc analysé les travaux [Pastor and Gay, 1995], [Isar and Naforniță, 1998], [Leporini and Pesquet, 1999] et mis en évidence le fait suivant. La  $M$ -TPOD est en effet beaucoup plus complexe que la TOD en ce sens où la plupart des *chemins* qui caractérisent l'*arbre* d'une  $M$ -TPOD, le paramètre de décalage  $n$  est une fonction non constante du niveau de décomposition  $j$ , fonction qui, de surcroît, ne peut être majorée par une constante indépendante de  $j$ . Plus précisément, on peut montrer que lorsque le paramètre de décalage  $n$  est supposé constant ou, variable mais borné par une constante, alors seul le chemin des coefficients d'approximation est considéré. Il résulte de cette analyse que l'outil commode pouvant être utilisé si  $n$  était constant (outil qui se trouve être le théorème de la convergence dominée de Lebesgue, voir [Leporini and Pesquet, 1999]) n'est pas applicable dans le contexte général de la  $M$ -TPOD.

Plus précisément, en utilisant les filtres d'analyse de Shannon et en adoptant une description des chemins de l'arbre de la TPOD sous forme de séquence d'éléments à valeur dans un alphabet binaire (dans une démarche similaire à celle donnée en [Pastor and Gay, 1995]), on peut montrer que la TPOD d'un processus stationnaire au sens large est telle que les fonctions d'autocorrélation asymptotiques sont différentes les unes des autres en fonction du chemin suivi : asymptotiquement, le processus obtenu dans chaque sous-bande est décorrelé, mais sa variance dépend du chemin choisi et peut être calculée en considérant la séquence binaire associée à ce chemin [Atto, Pastor, and Isar, 2007].

D'autre part, puisque certaines familles de filtres d'ondelettes (Daubechies, Battle-Lemarié) convergent vers les filtres idéaux de Shannon lorsque la régularité des filtres de ces familles augmente, on peut déduire, pour ces familles de filtres, un résultat analogue à celui obtenu pour la TPOD de Shannon, pourvu qu'on exploite également la convergence des filtres de ces familles vers les filtres de Shannon. Une telle analyse dépend donc à la fois des filtres d'analyse et du niveau de décomposition ([Atto, Pastor, and Isar, 2007]).

Plus généralement, [Atto and Pastor, 2008b] étend les résultats de [Atto, Pastor, and Isar, 2007] au cas de la  $M$ -TPOD, et adresse les distributions limites des coefficients obtenus dans les sous-bandes de la  $M$ -TPOD, par décomposition un processus aléatoire stationnaire au sens

strict. La convergence est étudiée par séquence de coefficients intra-bandes représentés sous forme de processus aléatoire discret. Il est ainsi montré que le processus aléatoire représentant les coefficients intra-bandes de la  $M$ -TPOD d'un processus strictement stationnaire tend vers un processus blanc et Gaussien, de variance égale à la valeur de la densité spectrale du processus décomposé en une certaine fréquence, cette fréquence étant elle-même étroitement liée au chemin parcouru dans l'arbre des paquets d'ondelettes. Cette convergence est explicitée dans le Chapitre 3 et concerne les filtres d'ondelettes ci-dessus décrits [Atto, Pastor, and Isar, 2007].

L'organisation de cette partie, consacrée aux propriétés statistiques de la  $M$ -TPOD, est la suivante. Le Chapitre 1 introduit les principes de bases de la  $M$ -TPOD à partir du *Splitting Lemma*. Le Chapitre 2 est dédié à l'étude de la corrélation statistique des coefficients intra-bandes de la  $M$ -TPOD. Le processus décomposé sera supposé stationnaire au sens large dans ce chapitre. Le Chapitre 3 est consacré à l'étude de la distribution des coefficients intra-bandes de la  $M$ -TPOD. Le processus décomposé sera supposé stationnaire au sens strict dans ce chapitre. Les Chapitres 2 et 3 présentent également quelques tests expérimentaux dans le but d'illustrer les résultats que l'on peut espérer atteindre en pratique, avec des familles de filtres standards, et lorsque le niveau de décomposition est fini.

# CHAPITRE 1 Généralités

Ce chapitre est consacré aux rappels, définitions et notations qui seront utilisées dans cette partie du document. Il propose une synthèse de la théorie des ondelettes et analyse certains fondements de cette théorie. Nous considérons l'espace de Hilbert  $L^2(\mathbb{R})$  compte tenu du fait qu'en traitement de signaux (mono comme bidimensionnels), la plupart des signaux considérés sont d'énergie finie. Quelques rappels d'analyse, ainsi que les détails concernant les représentations dans les espaces de Hilbert sont donnés dans l'annexe A. Le lecteur peut se référer à [Daubechies, 1992], [Antoniadis and Oppenheim, 1995], [Härdle, Kerkycharian, Picard, and Tsybakov, 1998], [Mallat, 1999] et [Misiti, Misiti, Oppenheim, and Poggi, 2003] pour de plus amples détails. Les TOD et  $M$ -TPOD sont des représentations Hilbertiennes pour lesquelles, la base de Hilbert sur laquelle s'effectue la projection est générée à partir d'une seule fonction appartenant à  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  et appelée *fonction d'ondelette* (voir annexe A.2.3).

Pour simplifier la lecture de ce chapitre, nous introduisons d'abord le principe de la  $M$ -TPOD car la TPOD classique en est un cas particulier, le cas qui correspond à  $M = 2$ . Ensuite, la TOD est décrite comme une application particulière du principe de décomposition de la TPOD. Nous précisons les spécificités de ces transformées au fur et à mesure. D'autre part, nous présentons uniquement les principes de ces transformées dans le cas où les bases d'ondelettes sont des bases orthonormales.

## 1.1 Représentations en ondelettes et bancs de filtres à reconstruction parfaite

Une  $M$ -TPOD est une représentation qui s'implémente à l'aide d'un banc de filtres à reconstruction parfaite [Steffen, Heller, Gopinath, and Burrus, 1993]. Un banc de filtres  $H_m$ ,  $m = 0, 1, 2, \dots, M-1$ , avec

$$H_m(\omega) = \frac{1}{\sqrt{M}} \sum_{\ell \in \mathbb{Z}} h_m[\ell] e^{-i\ell\omega}, \quad (1.1)$$

est dit à reconstruction parfaite s'il satisfait les propriétés suivantes.

1.  $h_m$  (réponse impulsionnelle de  $H_m$ ) est un élément de  $\ell^2(\mathbb{Z})$  pour tout  $m = 0, 1, 2, \dots, M-1$ .

## 2. La matrice

$$\begin{pmatrix} H_0(\omega) & H_1(\omega) & \cdots & H_{M-1}(\omega) \\ H_0(\omega + \frac{\pi}{M}) & H_1(\omega + \frac{\pi}{M}) & \cdots & H_{M-1}(\omega + \frac{\pi}{M}) \\ \vdots & \vdots & \ddots & \vdots \\ H_0(\omega + \frac{(M-1)\pi}{M}) & H_1(\omega + \frac{(M-1)\pi}{M}) & \cdots & H_{M-1}(\omega + \frac{(M-1)\pi}{M}) \end{pmatrix},$$

est unitaire pour tout nombre réel  $\omega$ .

L'unitarité de cette matrice implique  $|H_m(\omega)| \leq 1$  pour tout  $m = 0, 1, \dots, M-1$  et tout  $\omega \in \mathbb{R}$ .

Une famille  $\psi_m$ ,  $m = 1, \dots, M-1$  de fonctions d'ondelettes sera dite associée à une fonction d'échelle  $\phi$  s'il existe un banc de filtres à reconstruction parfaite  $H_m$ ,  $m = 0, 1, \dots, M-1$  tel que

$$\mathcal{F}\phi(\omega) = H_0(\omega/M)\mathcal{F}\phi(\omega/M), \quad (1.2)$$

et, pour tout  $m = 1, \dots, M-1$ ,

$$\mathcal{F}\psi_m(\omega) = H_m(\omega/M)\mathcal{F}\phi(\omega/M). \quad (1.3)$$

On dira alors du banc de filtre  $H_m$ ,  $m = 0, 1, \dots, M-1$ , qu'il est associé à une  $M$ -TPOD.

Notons que la donnée d'une fonction d'échelle admissible permet de définir le filtre d'échelle  $H_0$  à partir de l'Eq. (1.2). Connaissant  $H_0$ , on peut calculer les filtres  $H_m$ ,  $m = 1, 2, \dots, M-1$  compte tenu de la condition de reconstruction parfaite. On peut ainsi calculer la famille de fonctions d'ondelettes associée à cette fonction d'échelle grâce à l'Eq. (1.3). On dit que le filtre  $H_0$  est associé à la fonction d'échelle  $\phi$ . Ce filtre est d'ailleurs appelé (couramment) *filtre d'échelle*. Notons qu'une fonction d'échelle  $\phi$  admissible peut être calculée par normalisation d'une fonction  $\phi_0$  satisfaisant aux conditions suivantes (voir [Daubechies, 1992] pour le cas  $M=2$ ) :

- (1) Les fonctions  $\phi_0$  et  $\mathcal{F}\phi_0$  sont à décroissance suffisamment rapide.
- (2)  $\int \phi_0(t) dt \neq 0$ .
- (3) Il existe  $x = (x_\ell)_\ell \in \ell^2(\mathbb{Z})$  tel que  $\phi_0(t) = \sum_\ell x_\ell \phi_0(Mt - \ell)$ .
- (4) Il existe  $0 < \alpha \leq \beta < +\infty$  tel que  $\alpha \leq \sum_{\ell \in \mathbb{Z}} |\mathcal{F}\phi_0(\omega + M\pi\ell)|^2 \leq \beta$ .

Étant donné une fonction d'échelle  $\phi$  et les fonctions d'ondelettes associées  $\psi_m$ ,  $m = 1, 2, \dots, M-1$ , les coefficients filtres de la  $M$ -TPOD peuvent se calculer par les relations  $h_0[\ell] = \frac{1}{\sqrt{M}} \langle \phi(t), \phi(Mt - \ell) \rangle$ , et, pour  $m = 1, \dots, M-1$ ,  $h_m[\ell] = \frac{1}{\sqrt{M}} \langle \psi_m(t), \phi(Mt - \ell) \rangle$ .

## 1.2 $M$ -TPOD : Principe de la décomposition

Soit  $\Phi$ , une fonction telle que  $\{\tau_k \Phi : k \in \mathbb{Z}\}$  est un système orthonormal de fonctions de  $L^2(\mathbb{R})$ , où  $\tau_k \Phi : t \mapsto \Phi(t - k)$ . Soit  $\mathbf{U}$  l'adhérence (fermeture) du sous-espace engendré par ce système orthonormal.

Soit  $H_m$ ,  $m = 0, 1, 2, \dots, M-1$ , un banc de filtres à reconstruction parfaite associé à une  $M$ -TPOD. On définit une séquence d'élément de  $L^2(\mathbb{R})$  par

$$W_m(t) = \sqrt{M} \sum_{\ell \in \mathbb{Z}} h_m[\ell] \Phi(Mt - \ell) \quad (1.4)$$

et

$$W_{Mn+m}(t) = \sqrt{M} \sum_{\ell \in \mathbb{Z}} h_m[\ell] W_n(Mt - \ell) \quad (1.5)$$

pour  $n \in \mathbb{N}$  et  $m = 0, 1, 2, \dots, M-1$ . L'Eq. (1.4) peut encore s'écrire sous la forme

$$\mathcal{F}W_m(\omega) = H_m(\omega/M) \mathcal{F}\Phi(\omega/M). \quad (1.6)$$

De la même manière, on obtient de l'Eq. (1.5)

$$\mathcal{F}W_{Mn+m}(\omega) = H_m(\omega/M) \mathcal{F}W_n(\omega/M) \quad (1.7)$$

pour  $n \in \mathbb{N}$  et  $m = 0, 1, \dots, M-1$ .

Notons que la fonction  $\Phi$  dans l'Eq. (1.4) n'est pas nécessairement la fonction d'échelle associée au filtre  $H_0$ . Lorsque  $\Phi$  est la fonction d'échelle associée à  $H_0$ , nous avons  $W_0 = \Phi$ . Dans ce cas, les Eqs. (1.5) et (1.7) sont vraies pour  $n = 0$ .

Pour tout couple  $(j, n)$  d'entiers naturels, posons

$$W_{j,n}(t) = M^{-j/2} W_n(M^{-j} t). \quad (1.8)$$

D'après les Eqs. (1.1), (1.7) et (1.8), nous avons

$$\mathcal{F}W_{j+1,Mn+m}(\omega) = \sqrt{M} H_m(M^j \omega) \mathcal{F}W_{j,n}(\omega). \quad (1.9)$$

Les formules (1.6), (1.7) et (1.9) sont satisfaites presque partout [Daubechies, 1992]. Posons  $W_{j,n,k} = \tau_{M^j k} W_{j,n}$ , ce qui revient à

$$W_{j,n,k}(t) = W_{j,n}(t - M^j k) = M^{-j/2} W_n(M^{-j} t - k). \quad (1.10)$$

L'ensemble  $\{W_{j,n,k} : k \in \mathbb{Z}\}$  est un système orthonormal de  $L^2(\mathbb{R})$ . L'adhérence de l'espace fonctionnel engendré par  $\{W_{j,n,k} : k \in \mathbb{Z}\}$  sera appelé dans la suite espace de *paquets d'ondelettes*  $\mathbf{W}_{j,n}$ . Tout  $W_{j,n,k}$ , et donc, tout  $W_{j,n}$  puisque  $W_{j,n} = W_{j,n,0}$  et aussi tout  $W_n$  puisque  $W_n = W_{0,n,0}$ , est appelé par la suite fonction associée à un paquet d'ondelettes ou tout simplement *fonction d'ondelette*.

La M-TPOD de l'espace fonctionnel  $\mathbf{U}$  consiste à décomposer (*diviser*) l'espace  $\mathbf{U}$  en  $M$  sous-espaces orthogonaux :

$$\mathbf{U} = \bigoplus_{m=0}^{M-1} \mathbf{W}_{1,m}, \quad (1.11)$$

et à appliquer récursivement le même principe de division (Splitting lemma, voir [Daubechies, 1992] pour  $M=2$ ) :

$$\mathbf{W}_{j,n} = \bigoplus_{m=0}^{M-1} \mathbf{W}_{j+1,Mn+m}, \quad (1.12)$$

pour tout entier naturel  $j$  et tout  $n = 0, 1, 2, \dots, M^j - 1$ . Dans cette décomposition,  $\{W_{j+1,Mn+m,k} : k \in \mathbb{Z}\}$  est une base orthonormale de l'espace  $\mathbf{W}_{j+1,Mn+m}$ . Les arbres des décompositions en 2-TOPD et 3-TOPD sont donnés à titre d'exemple par les figures 1.1 et 1.2.

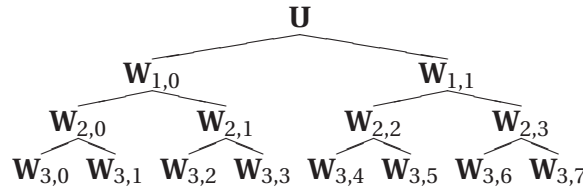


Figure 1.1 — Arbre de la décomposition en 2-TPOD de  $\mathbf{U}$ , pour 3 niveaux de décompositions ( $j = 3$ ).

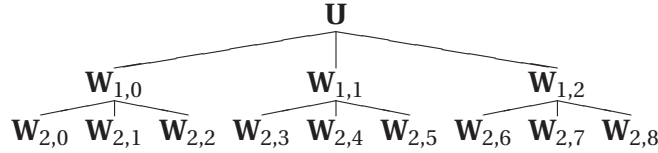


Figure 1.2 — Arbre de la décomposition en 3-TPOD de  $\mathbf{U}$ , pour 2 niveaux de décompositions ( $j = 3$ ).

D'après les Eqs. (1.11) et (1.12),  $\mathbf{U}$  peut être divisé en sommes orthogonales de paquets d'ondelettes. Ainsi, étant donné un élément  $f$  de  $\mathbf{U}$ , les coefficients de la projection de  $f$  sur un espace  $\mathbf{W}_{j,n}$  définit une séquence de *coefficients d'ondelettes*  $(c_{j,n}[k])_{k \in \mathbb{Z}}$ , avec

$$c_{j,n}[k] = \int_{\mathbb{R}} f(t) W_{j,n,k}(t) dt. \quad (1.13)$$

Les paquets d'ondelettes considérés par la suite requièrent le calcul de la  $M$ -TPOD par les Eqs. (1.11) et (1.12). Par conséquent, tout  $\mathbf{W}_{j,n}$  rencontré dans la suite est tel que  $n \in \{0, \dots, M^j - 1\}$ .

Comme signalé un peu plus haut, la 2-TOPD correspond à la TOPD classique. D'autre part, la TOD est obtenue à partir de la TPOD en considérant une décomposition de  $\mathbf{U}$  dans laquelle les espaces  $\mathbf{W}_{j,1}$  ne sont pas concernés par la division. Dans ce cas, on a

$$\mathbf{U} = \mathbf{W}_{1,0} \oplus \mathbf{W}_{1,1} = \mathbf{W}_{J,0} \bigoplus_{j=1}^J \mathbf{W}_{j,1}, \quad (1.14)$$

pour tout  $J \geq 1$ .

### 1.3 Représentation des chemins de l'arbre de décomposition de la $M$ -TPOD à l'aide séquences $M$ -aires

Un *chemin*  $\mathcal{P}$  de l'arbre de la  $M$ -TPOD est décrit par une séquence de sous-espaces fermés emboîtés :  $\mathcal{P} = (\mathbf{U}, \{\mathbf{W}_{j,n(j)}\}_{j \in \mathbb{N}})$ , où  $\mathbf{W}_{j,n(j)} \subset \mathbf{W}_{j-1,n(j-1)}$ . Par construction, chaque  $\mathbf{W}_{j,n(j)}$  est obtenu par décomposition récursive de  $\mathbf{U}$  à l'aide de séquences particulières de filtres  $(h_{m_\ell})_{\ell=1,2,\dots,j}$  où chaque  $m_\ell$  appartient à  $\{0, 1, \dots, M-1\}$ . Il en résulte que le paramètre de décalage est

$$n(j) = \sum_{\ell=1}^j m_\ell M^{j-\ell} \in \{0, 1, \dots, M^j - 1\} \quad (1.15)$$

pour tout niveau de décomposition  $j$ . Ainsi, le chemin  $\mathcal{P}$  peut être associé à une séquence  $M$ -aire  $(m_\ell)_{\ell \geq 1}$  d'éléments : pour tout  $\ell \geq 1$ ,  $m_\ell$  appartient à l'alphabet à  $M$  éléments  $\{0, 1, \dots, M-1\}$ .



1}. Cette séquence caractérise les paquets  $\{\mathbf{W}_{j,n(j)}\}_{j \in \mathbb{N}}$ , avec un paramètre de décalage  $n = n(j)$  donné pour tout niveau de décomposition  $j$  par l'Eq. (1.15). On a aussi la relation

$$n(j) = Mn(j-1) + m_j, \quad (1.16)$$

pour  $j \in \mathbb{N}$ , avec la convention  $n(0) = 0$ .

Réciproquement, étant donné une séquence  $M$ -aire  $\kappa = (m_\ell)_{\ell \in \mathbb{N}}$  pour laquelle chaque  $m_\ell$  est un élément de  $\{0, 1, \dots, M-1\}$ , la sous-séquence finie  $(m_\ell)_{\ell=1,2,\dots,j}$ , formée par les  $j$  premiers termes de  $\kappa$ , définit un unique et non négatif entier relatif  $n(j)$  de la forme de l'Eq. (1.15) et est donc associé à un unique paquet d'ondelettes localisé au nœud  $(j, n(j))$  de l'arbre de la décomposition par  $M$ -TPOD.

D'après ces remarques, on peut associer tout chemin de l'arbre de la  $M$ -TPOD à une unique séquence  $\kappa = (m_\ell)_{\ell \in \mathbb{N}}$  d'éléments de  $\{0, 1, \dots, M-1\}$  et *vice versa*. Dans ce qui suit, nous adopterons cette description des chemins par des séquences  $M$ -aires, lorsque cela est plus commode. Puisque le paramètre de décalage  $n$  dépend de  $j$  et  $\kappa$  via l'Eq. (1.15), la notation  $n = n_\kappa(j)$  sera utilisée dans la suite pour indiquer cette dépendance, lorsque cela est utile. Pour résumer cette partie, nous dirons que toute séquence  $M$ -aire  $\kappa$  associée à un chemin de l'arbre de la  $M$ -TPOD spécifie une unique séquence  $(\mathbf{W}_{j,n_\kappa(j)})_{j \in \mathbb{N}}$  de paquets d'ondelettes. Notons aussi que les espaces  $\mathbf{W}_{j,0}$ ,  $j \in \mathbb{N}$ , sont appelés *espaces d'approximations* du fait de l'utilisation du filtre  $h_0$  à chaque niveau de décomposition. Les espaces  $\mathbf{W}_{j,n}$ ,  $j \in \mathbb{N}$ ,  $n \in \{1, \dots, M^j - 1\}$ , sont appelés *espaces de détails*.

**Exemple 1** Soit la séquence

$$\kappa_0 = (0, 0, 0, 0, 0, 0, \dots).$$

Nous avons  $n_{\kappa_0}(j) = 0$  pour tout entier naturel  $j$ . Ainsi,  $\kappa_0$  est associé au chemin  $\mathcal{P}_{\kappa_0} = (\mathbf{U}, \{\mathbf{W}_{j,0}\}_{j \in \mathbb{N}})$  situé le plus à gauche dans l'arbre de la  $M$ -TPOD. Ce chemin est appelé *chemin des approximations* car il correspond aux espaces d'approximations  $\mathbf{W}_{j,0}$ ,  $j \in \mathbb{N}$ .

Considérons maintenant la séquence

$$\kappa_1 = (M-1, 0, 0, 0, 0, 0, \dots).$$

Pour cette séquence, nous avons  $n_{\kappa_1}(j) = (M-1)M^{j-1}$ . Cette séquence est donc associée au chemin  $\mathcal{P}_{\kappa_1} = (\mathbf{U}, \{\mathbf{W}_{j,(M-1)M^{j-1}}\}_{j \in \mathbb{N}})$  obtenu par utilisation du filtre  $h_{M-1}$  au premier niveau de décomposition et le filtre  $h_0$  pour tout niveau de décomposition  $j \geq 2$ .

Nous concluons cette section par le lemme suivant, dont l'intérêt sera mis en évidence par la suite.

**Lemme 1** Pour  $n = n(j)$  donné par l'Eq. (1.15), on a

$$\mathcal{F}W_n(\omega) = \left[ \prod_{\ell=1}^j H_{m_\ell} \left( \frac{\omega}{M^{j+1-\ell}} \right) \right] \mathcal{F}\Phi \left( \frac{\omega}{M^j} \right). \quad (1.17)$$



**Preuve :** Des Eqs. (1.7) et (1.16), on a

$$\mathcal{F}W_{n(j)}(\omega) = \mathcal{F}W_{Mn(j-1)+m_j}(\omega) = H_{m_j}\left(\frac{\omega}{M}\right)\mathcal{F}W_{n(j-1)}\left(\frac{\omega}{M}\right).$$

Par récurrence, on obtient

$$\mathcal{F}W_{n(j)}(\omega) = H_{m_j}\left(\frac{\omega}{M}\right)H_{m_{j-1}}\left(\frac{\omega}{M^2}\right)\dots H_{m_2}\left(\frac{\omega}{M^{j-1}}\right)\mathcal{F}W_{m_1}\left(\frac{\omega}{M^{j-1}}\right),$$

puisque  $n(1) = m_1$ . Il suffit alors d'appliquer l'Eq. (1.6) pour conclure. ■

## 1.4 La $M$ -TPOD de Shannon et l'espace de Paley-Wiener des fonctions à bandes limitées

Les filtres de la  $M$ -TPOD de Shannon sont des filtres idéaux passe-bas, passe-haut, et passe-bande. Ils sont notés  $h_m^S$  pour  $m = 0, 1, \dots, M-1$ . La transformée de Fourier discrète de  $h_m^S$  est donnée par

$$H_m^S(\omega) = \sum_{\ell \in \mathbb{Z}} \mathbb{1}_{\Delta_m}(\omega - 2\pi\ell), \quad (1.18)$$

où  $\Delta_m = \left[-\frac{(m+1)\pi}{M}, -\frac{m\pi}{M}\right] \cup \left[\frac{m\pi}{M}, \frac{(m+1)\pi}{M}\right]$ . La fonction d'échelle  $\Phi^S$  associée à ces filtres est définie pour tout  $t \in \mathbb{R}$  par  $\Phi^S(t) = \text{sinc}(t) = \sin(\pi t)/\pi t$ , avec  $\Phi^S(0) = 1$ . La transformée de Fourier de cette fonction d'échelle est

$$\mathcal{F}\Phi^S = \mathbb{1}_{[-\pi, \pi]}, \quad (1.19)$$

où  $\mathbb{1}_K$  est la fonction indicatrice de l'intervalle  $K$  :  $\mathbb{1}_K(x) = 1$  si  $x \in K$  et  $\mathbb{1}_K(x) = 0$ , sinon. L'adhérence  $\mathbf{U}^S$  de l'espace engendré par le système orthonormal  $\{\tau_k\Phi^S : k \in \mathbb{Z}\}$  est alors l'espace de Paley-Wiener (PW) des éléments de  $L^2(\mathbb{R})$  qui sont à bande limitée en ce sens où leurs transformées de Fourier sont à support dans  $[-\pi, \pi]$ . Tout élément de cet espace satisfait au théorème d'échantillonnage de Shannon. Par conséquent, lorsque la  $M$ -TPOD concerne l'espace  $\mathbf{U}^S$ , les données en entrée de la décomposition de tout élément  $f$  de  $\mathbf{U}^S$  sont les échantillons  $\{f[k]\}_{k \in \mathbb{Z}}$  de  $f$ .

L'espace  $\mathbf{U}^S$  jouera un rôle primordial dans ce qui suit, pour la raison suivante. Soit  $X$  un processus aléatoire WSS et à bande limitée (dont le spectre est à support dans  $[-\pi, \pi]$ ). Alors on a (voir [Atto, Pastor, and Isar, 2007, Appendix D])

$$X[k] = \int_{\mathbb{R}} X(t)\Phi^S(t-k)dt, \quad (1.20)$$

de sorte que l'espace  $\mathbf{U}^S$  est l'espace naturel de représentation d'un tel processus. Une  $M$ -TPOD de  $X$  peut donc être initialisée par les échantillons  $X[k]$ ,  $k \in \mathbb{Z}$ .

Considérons la  $M$ -TPOD de Shannon de l'espace  $\mathbf{U}^S$ . Les fonctions d'ondelettes  $W_{j,n}^S$  de cette  $M$ -TPOD peuvent être calculées selon les Eqs. (1.4), (1.5) et (1.8) en posant  $\Phi = \Phi^S$  et  $h_m = h_m^S$ ,  $m = 0, 1, \dots, M-1$ . La transformée de Fourier de la fonction  $W_{j,n}^S$  est donnée par le lemme suivant.

**Proposition 1** Pour tout  $j \in \{0, 1, 2, \dots\}$  et tout  $n \in \{0, \dots, M^j - 1\}$ , on a

$$\mathcal{F}W_{j,n}^S = M^{j/2} \mathbb{1}_{\Delta_{j,G(n)}}, \quad (1.21)$$

où

$$\Delta_{j,k} = \left[ -\frac{(k+1)\pi}{M^j}, -\frac{k\pi}{M^j} \right] \cup \left[ \frac{k\pi}{M^j}, \frac{(k+1)\pi}{M^j} \right] \quad (1.22)$$

pour tout  $k \in \mathbb{Z}^+$ , et  $G$  est l'application définie récursivement par  $G(0) = 0$  et

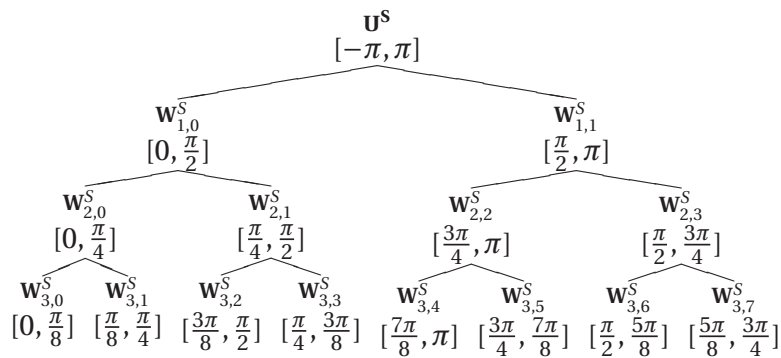
$$G(M\ell + m) = \begin{cases} MG(\ell) + m & \text{si } G(\ell) \text{ est paire,} \\ MG(\ell) - m + M - 1 & \text{si } G(\ell) \text{ est impaire,} \end{cases} \quad (1.23)$$

pour  $m = 0, 1, \dots, M-1$  et  $\ell = 0, 1, 2, \dots$

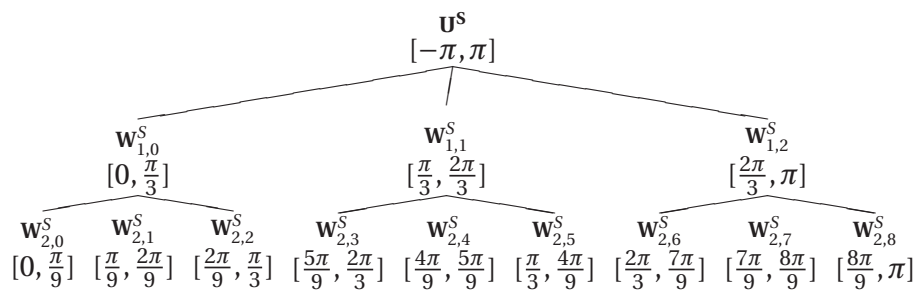
**Preuve :** [Atto and Pastor, 2008b]. ■

Le support de  $\mathcal{F}W_{j,n}^S$  est  $\Delta_{j,p}$ , où  $p = G(n)$ . L'ensemble  $\Delta_{j,p}$  est appelé sous-bande. La  $M$ -TPOD de Shannon de l'espace  $\mathbf{U}^S$  est le codage idéal en sous-bandes. Par la suite, on parle de coefficients intra-bande (ou de sous-bande) lorsqu'on considère la projection sur un espace  $\mathbf{W}_{j,n}$  quelconque. Notons aussi que la restriction de  $G$  à l'ensemble  $\{0, 1, \dots, M^j - 1\}$  est une permutation de cet ensemble. Cette permutation induit un ré-arrangement fréquentiel des paquets d'ondelettes de Shannon,  $\mathcal{F}W_{j,n}^S$ ,  $n = 0, 1, \dots, M^j - 1$ . Ce ré-arrangement fréquentiel est donné pour les 2-TPOD et 3-TPOD de Shannon par les figures 1.3 et 1.4 respectivement. Dans ces figures, la partie positive  $\Delta_{j,p}^+$  du support de la fonction  $\mathcal{F}W_{j,n}^S$ , est donné pour chaque niveau de décomposition  $j$  considéré et pour tout  $n = 0, 1, \dots, M^j - 1$ . L'ensemble  $\Delta_{j,k}^+$  est donné pour  $(j, k) \in \mathbb{Z}^+ \times \mathbb{Z}^+$  par

$$\Delta_{j,k}^+ = \left[ \frac{k\pi}{M^j}, \frac{(k+1)\pi}{M^j} \right]. \quad (1.24)$$



**Figure 1.3** — Ordre fréquentiel des paquets de l'arbre de la 2-TPOD de Shannon. La partie positive du support de  $\mathcal{F}W_{j,n}^S$  est indiquée à côté de  $\mathbf{W}_{j,n}^S$ .



**Figure 1.4** — Ordre fréquentiel des paquets de l'arbre de la 3-TPOD de Shannon. La partie positive du support de  $\mathcal{F}W_{j,n}^S$  est indiquée à côté de  $\mathbf{W}_{j,n}^S$ .

## Analyse asymptotique de la corrélation intra-bande de la $M$ -TPOD

### 2.1 $M$ -TPOD des processus aléatoires stationnaires au sens large - Résultats préliminaires

Soit  $X$ , un processus réel, centré, du second ordre et continu en moyenne quadratique. La fonction d'autocorrélation de  $X$ , notée  $R$ , est définie par

$$R(t, s) = \mathbb{E}[X(t)X(s)].$$

La projection de  $X$  sur  $\mathbf{W}_{j,n}$  résulte en une séquence de variables aléatoires, les *coefficients* du processus aléatoire  $X$  :

$$c_{j,n}[k] = \int_{\mathbb{R}} X(t) W_{j,n,k}(t) dt, \quad k \in \mathbb{Z}, \quad (2.1)$$

lorsque l'intégrale

$$\iint_{\mathbb{R}^2} R(t, s) W_{j,n,k}(t) W_{j,n,\ell}(s) dt ds \quad (2.2)$$

existe [Houdré, 1994], ce qui sera admis par la suite puisque les fonctions d'ondelettes communément utilisées sont à support compact ou à décroissance suffisamment rapide. La séquence donnée par l'Eq. (2.1) définit le processus aléatoire discret  $c_{j,n} = (c_{j,n}[k])_{k \in \mathbb{Z}}$  des *coefficients* d'ondelettes de  $X$  au niveau de décomposition  $j$  et pour un paramètre de décalage  $n$  donné. Soit  $R_{j,n}$  la fonction d'autocorrélation du processus  $c_{j,n}$ . On a

$$R_{j,n}[k, \ell] = \mathbb{E}[c_{j,n}[k]c_{j,n}[\ell]] = \iint_{\mathbb{R}^2} R(t, s) W_{j,n,k}(t) W_{j,n,\ell}(s) dt ds. \quad (2.3)$$

Si  $X$  est stationnaire au sens large, on peut écrire (par abus de langage)  $R(t, s) = R(t - s)$ . Ainsi, et à partir de l'Eq. (2.3), on obtient

$$R_{j,n}[k, \ell] = \iint_{\mathbb{R}^2} R(t) W_{j,n,k}(t + s) W_{j,n,\ell}(s) dt ds. \quad (2.4)$$

Dans ce qui suit, et ce jusqu'à la fin de ce chapitre, nous supposons que  $X$  admet un spectre (densité spectrale de puissance)  $\gamma$  qui est, par définition, la transformée de Fourier de  $R$ . En utilisant le théorème de Fubini et l'égalité de Parseval, on montre (voir [Atto, Pastor, and Isar, 2007, Appendix C], [Atto and Pastor, 2008b]) à partir des Eqs. (1.10) et (2.4) que la fonction d'autocorrélation  $c_{j,n}$  s'écrit :

$$R_{j,n}[k, \ell] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma\left(\frac{\omega}{M^j}\right) |\mathcal{F}W_n(\omega)|^2 e^{i(k-\ell)\omega} d\omega. \quad (2.5)$$

Ainsi,  $c_{j,n}$  est un processus aléatoire stationnaire au sens large. Pour tout  $k, \ell \in \mathbb{Z}$ , la valeur  $R_{j,n}[k, \ell]$  de la fonction d'autocorrélation du processus aléatoire discret  $c_{j,n}$  est  $R_{j,n}[k - \ell]$ , avec

$$R_{j,n}[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma\left(\frac{\omega}{M^j}\right) |\mathcal{F}W_n(\omega)|^2 e^{ik\omega} d\omega. \quad (2.6)$$

Supposons que la fonction  $\gamma \in L^\infty(\mathbb{R})$  et est continue en 0. Il en résulte les conséquences suivantes. D'abord, l'intégrande dans l'Eq. (2.6) est intégrable puisque sa valeur absolue est majorée par  $\|\gamma\|_\infty |\mathcal{F}W_n(\cdot)|^2$ , dont l'intégrale vaut  $\|\gamma\|_\infty$ ; ensuite, la limite lorsque  $j$  tend vers  $\infty$  de  $\gamma(\frac{\omega}{M^j})$  est  $\gamma(0)$ . Ainsi, pour tout entier naturel  $n$  fixé, il découle du théorème de la convergence dominée de Lebesgue appliqué à l'Eq. (2.6) que

$$\lim_{j \rightarrow +\infty} R_{j,n}[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma(0) |\mathcal{F}W_n(\omega)|^2 e^{ik\omega} d\omega = \gamma(0) \delta[k], \quad (2.7)$$

où  $\delta[\cdot]$  désigne le symbole de Kronecker : pour tout  $k \in \mathbb{Z}$ ,

$$\delta[k] = \begin{cases} 1 & \text{si } k = 0, \\ 0 & \text{si } k \neq 0, \end{cases}$$

résultat qui se déduit de [Leporini and Pesquet, 1999, Corollary 5].

Soit alors  $\mathcal{P} = (\mathbf{U}, \{\mathbf{W}_{j, n_\kappa(j)}\}_{j \in \mathbb{N}})$ , un chemin de l'arbre de la  $M$ -TPOD, avec  $\kappa = (m_\ell)_{\ell \in \mathbb{N}}$ , une séquence d'éléments de  $\{0, 1, \dots, M-1\}$ . Pour chaque niveau de décomposition  $j$ , le paramètre de décalage  $n = n_\kappa(j)$  dépend de  $j$ . On a deux cas de figure :

- (i)  $n_\kappa$  est constant avec  $j$ , ou majoré par une constante qui ne dépend pas de  $j$ . Dans ce cas, on obtient de l'Eq. (1.16) que  $\kappa = \kappa_0$  est la séquence nulle donnée dans l'exemple 1. Le paramètre de décalage  $n_\kappa(j)$  vaut donc 0 à chaque niveau de décomposition et la  $M$ -TPOD associée au chemin  $\mathcal{P} = \mathcal{P}_{\kappa_0}$  consiste en une séquence infinie de filtrage passe-bas. La décorrélation est garantie dans ce cas par l'Eq. (2.7) (voir aussi [Leporini and Pesquet, 1999, Corollary 5]).
- (ii)  $n_\kappa$  n'est pas constant avec  $j$  (et n'est pas majorable par une constante indépendante de  $j$ ). Un exemple est la séquence  $\kappa_1$  donnée dans l'exemple 1, ou encore la séquence  $\kappa = (1, 1, \dots)$  pour laquelle  $n_\kappa(j) = M^j - 1$  de sorte que les nœuds  $(j, M^j - 1)$  sont ceux situés à l'extrémité droite de l'arbre de la  $M$ -TPOD (voir figures 1.1 et 1.2). Dans ce cas où  $n_\kappa$  est une fonction dépendant explicitement de  $j$ , la décorrélation asymptotique des coefficients de la  $M$ -TPOD aux nœuds  $(j, n_\kappa(j))$ , lorsque  $j$  tend vers  $\infty$  ne peut pas être obtenue par application directe du théorème de la convergence dominée à partir des Eqs. (2.6) et (2.7).

Pour étudier le comportement des fonctions d'autocorrélation aux nœuds  $(j, n_\kappa(j))$  pour un chemin  $\kappa \neq \kappa_0$  donné, il est plus commode d'écrire les expressions des fonctions d'autocorrélation sous la forme

$$R_{j,n}[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma(\omega) |\mathcal{F}W_{j,n}(\omega)|^2 e^{iM^j k \omega} d\omega. \quad (2.8)$$

Cette expression provient de l'Eq. (2.6) après un changement de variable et en tenant compte du fait que

$$\mathcal{F}W_{j,n}(\omega) = M^{j/2} \mathcal{F}W_n(M^j \omega), \quad (2.9)$$

qui est déduit de l'Eq. (1.8).

Dans les Sections 2.2 et 2.3 suivantes, l'analyse du comportement asymptotique des fonctions d'autocorrélations  $(R_{j,n})_{j \geq 1}$  est réalisée lorsque la  $M$ -TPOD concerne les filtres de Shannon ou certaines familles particulières de filtres. Nous étudierons d'abord le cas où la  $M$ -TPOD est réalisée avec les filtres de Shannon présentés dans la Section 1.4. Le résultat obtenu dans ce contexte est le Théorème 1. Ensuite, nous considérons le cas de la  $M$ -TPOD réalisée avec des filtres qui approximent, dans le sens spécifié ci-après, ceux de Shannon. Le résultat obtenu dans ce cas est le Théorème 2. Les Théorèmes 1 et 2 concernent des processus stationnaires au sens large et à bandes limitées. Par conséquent, l'espace de décomposition considéré est l'espace  $\mathbf{U}^S$  de PW puisqu'il est l'espace de représentation naturel de tels processus (voir l'Eq. (1.20)).

## 2.2 Autocorrélations asymptotiques de la $M$ -TPOD de Shannon

### 2.2.1 Convergence des fonctions d'autocorrélations

Soit  $\kappa = (m_\ell)_{\ell \in \mathbb{N}}$  une séquence  $M$ -aire d'éléments de  $\{0, 1, \dots, M-1\}$ . Considérons la  $M$ -TPOD de Shannon qui est, par définition, la  $M$ -TPOD associée aux filtres  $(h_m^S)_{m=0,1,\dots,M-1}$ . Soit  $\mathcal{P}_\kappa = (\mathbf{U}^S, \{\mathbf{W}_{j,n_\kappa(j)}^S\}_{j \in \mathbb{N}})$ , le chemin associé à  $\kappa$  dans l'arbre de la  $M$ -TPOD. D'après la Proposition 1, le support de  $W_{j,n_\kappa(j)}^S$  est  $\Delta_{j,p_\kappa(j)}$ , avec  $p_\kappa(j) = G(n_\kappa(j))$ . Pour  $j \in \mathbb{N}$ , les intervalles  $\Delta_{j,p_\kappa(j)}^+$  sont fermés, emboîtés et leur diamètre tend vers 0. On en déduit que leur intersection est réduite à un seul point  $a(\kappa)$  donné par :

$$a(\kappa) = \lim_{j \rightarrow +\infty} \frac{p_\kappa(j)\pi}{M^j}. \quad (2.10)$$

Notons qu'on obtient ce résultat de manière équivalente, en remarquant que les bornes des intervalles  $\Delta_{j,p_\kappa(j)}^+$  définissent deux suites de Cauchy  $(\frac{p_\kappa(j)\pi}{M^j})_j$  et  $(\frac{(p_\kappa(j)+1)\pi}{M^j})_j$  admettant la même limite, donc  $a(\kappa)$  donné par l'Eq. (2.10). En effet, selon l'Eq. (1.12), on a  $n_\kappa(j+1) = Mn_\kappa(j) + m$ , pour un certain  $m \in \{0, 1, \dots, M-1\}$ . Compte tenu de l'Eq. (1.23), on a aussi  $p_\kappa(j) = Mp_\kappa(j) + m'$ , pour un certain  $m' \in \{0, 1, \dots, M-1\}$ . On obtient ainsi

$$\begin{aligned} p_\kappa(j+q) &= Mp_\kappa(j+q-1) + m'_{j+q} \\ &= M^q p_\kappa(j) + M^{q-1} m'_{j+1} + \dots + M^1 m'_{j+q-1} + m'_{j+q}, \end{aligned}$$

et puisque chaque  $m'_{j+\ell} \in \{0, 1, \dots, M-1\}$  pour  $\ell = 1, 2, \dots, q$ , on a donc,

$$\frac{p_\kappa(j+q)\pi}{M^{j+q}} - \frac{p_\kappa(j)\pi}{M^j} = \frac{1}{M^j} \sum_{\ell=1}^q \frac{m'_{\ell+j}}{M^\ell} \leq \frac{1}{M^j} - \frac{1}{M^{j+q}} < \frac{1}{M^j},$$

d'où le résultat.

Soit  $X$ , un processus du second ordre, stationnaire au sens large, continu en moyenne quadratique et de spectre  $\gamma$ . La fonction d'autocorrélation  $R_{j,n}^S$  résultant de la projection de  $X$  sur  $\mathbf{W}_{j,n}^S$  dérive de l'Eq. (2.8) et est donnée par

$$R_{j,n}^S[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma(\omega) |\mathcal{F} W_{j,n}^S(\omega)|^2 e^{iM^j k \omega} d\omega. \quad (2.11)$$

Des Eqs. (1.21) et (2.11), et compte tenu du fait que  $\gamma$  est paire, en tant que transformée de Fourier de la fonction paire  $R$ , on obtient

$$R_{j,n}^S[k] = \frac{M^j}{\pi} \int_{\Delta_{j,p}^+} \gamma(\omega) \cos(M^j k \omega) d\omega, \quad (2.12)$$

où  $\Delta_{j,p}^+$  est donné par l'Eq. (1.24) et  $p = G(n)$ .

**Théorème 1** *Soit  $X$ , un processus centré, du second ordre, stationnaire au sens large, et continu en moyenne quadratique. Supposons que  $X$  ait un spectre  $\gamma \in L^\infty(\mathbb{R})$  dont le support est contenu dans  $[-\pi, \pi]$ . Soit  $\kappa = (m_\ell)_{\ell \in \mathbb{N}}$  une séquence d'éléments à valeurs dans  $\{0, 1, \dots, M-1\}$  et  $\mathcal{P}_\kappa = (\mathbf{U}^S, \{\mathbf{W}_{j,n_\kappa(j)}^S\}_{j \in \mathbb{N}})$  le chemin associé à  $\kappa$  dans l'arbre de la  $M$ -TPOD.*

*Si le spectre  $\gamma$  est continu au point  $a(\kappa)$ , alors*

$$\lim_{j \rightarrow +\infty} R_{j,n_\kappa(j)}^S[k] = \gamma(a(\kappa)) \delta[k] \quad (2.13)$$

*uniformément en  $k \in \mathbb{Z}$ , où  $R_{j,n_\kappa(j)}^S$  désigne la fonction d'autocorrélation des coefficients obtenus par projection de  $X$  sur  $\mathbf{W}_{j,n_\kappa(j)}^S$ .*

**Preuve :** [Atto and Pastor, 2008b]. ■

Le Théorème 1 montre que les coefficients du processus  $X$  par projection sur  $(\mathbf{W}_{j,n_\kappa(j)})_j$  tendent à être décorrélés lorsque  $j$  tend vers l'infini. Ce résultat illustre la dépendance du processus de décorrélation avec le chemin spécifié, donc en fonction de la succession de filtrages effectuée.

**Exemple 2** *Dans le cas de la TPOD ( $M = 2$ ), les séquences  $\kappa_0$  et  $\kappa_1$  données à l'exemple 1 s'écrivent*

$$\kappa_0 = (0, 0, 0, 0, 0, 0, \dots), \quad \kappa_1 = (1, 0, 0, 0, 0, 0, \dots).$$

Nous considérons également les séquences

$$\kappa_2 = (0, 1, 0, 0, 0, 0, \dots), \quad \kappa_3 = (0, 0, 1, 0, 0, 0, \dots).$$

Les termes généraux des séquences  $\kappa_q$ ,  $q = 0, 1, 2$  et  $3$ , sont  $\delta[q - \ell]$  pour tout entier naturel  $\ell$ . De l'Eq. (1.15), on obtient, pour  $q = 1, 2, 3$ , que  $n_{\kappa_q}(j) = 0$  lorsque  $j = 1, 2, \dots, q - 1$ , et  $n_{\kappa_q}(j) = 2^{j-q}$  lorsque  $j = q, q + 1, \dots$ . Les séquences  $\kappa_q$ ,  $q = 1, 2$  et  $3$  sont donc telles que  $p_{\kappa_q}(j) = G(n_{\kappa_q}(j)) = 2^{j-q+1} - 1$  dès que  $j \geq q$ . On a ainsi  $\lim_{j \rightarrow \infty} p_{\kappa_q}(j)/2^j = 1/2^{q-1}$  pour  $q = 1, 2, 3$ . Par conséquent  $a(\kappa_q) = \pi/2^{q-1}$  pour  $q = 1, 2, 3$ . D'autre part, la séquence  $\kappa_0$  est telle que  $p_{\kappa_0}(j) = G(n_{\kappa_0}(j)) = 0$  pour tout  $j \in \mathbb{N}$ . ainsi  $a(\kappa_0) = 0$ . On conclut donc que pour  $q = 1, 2$  et  $3$ ,

$$\lim_{j \rightarrow +\infty} R_{j, n_{\kappa_q}}^S[k] = \gamma(\pi/2^{q-1})\delta[k], \quad (2.14)$$

tandis que

$$\lim_{j \rightarrow +\infty} R_{j, n_{\kappa_0}}^S[k] = \gamma(0)\delta[k]. \quad (2.15)$$

### 2.2.2 Vitesse de convergence du processus de décorrélation

On s'intéresse à la vitesse de convergence du processus de décorrélation de la  $M$ -TPOD de Shannon, c'est à dire à la vitesse de convergence vers la fonction d'autocorrélation limite. Les résultats donnés dans cette section sont une généralisation en bande  $M$  de la caractérisation obtenue pour  $M = 2$  dans [Atto, Pastor, and Isar, 2007].

Soit  $\omega_0 > 0$  et posons

$$\gamma(\omega) = \left(1 - \frac{|\omega|}{\omega_0}\right) \chi_{[-\pi, \pi] \cap [-\omega_0, \omega_0]}(\omega). \quad (2.16)$$

Supposons que  $\gamma$  représente le spectre d'un certain processus stationnaire au sens large. On a

$$R_{j, n}^S[k] = \frac{M^j}{\pi} \int_{\Delta_{j, p}^+ \cap [0, \omega_0]} \left(1 - \frac{\omega}{\omega_0}\right) \cos(M^j k \omega) d\omega. \quad (2.17)$$

On distingue deux cas :

(a) Si  $\omega_0 > \pi$ ,  $\Delta_{j, p}^+ \cap [0, \omega_0] = \Delta_{j, p}^+$ , et on obtient

$$R_{j, n}^S[k] = \begin{cases} 1 - \frac{\pi}{\omega_0} \frac{p}{M^j} - \frac{\pi}{\omega_0} \frac{1}{2M^j} & \text{if } k = 0 \\ \frac{(-1)^{kp} - (-1)^{k(p+1)}}{\pi \omega_0 k^2 M^j} & \text{if } k \neq 0, \end{cases} \quad (2.18)$$

d'où l'on conclut que,

$$\left| R_{j, n}^S[k] - \gamma\left(\frac{\pi p}{M^j}\right) \delta[k] \right| \leq \frac{\pi}{\omega_0} \frac{1}{2M^j}. \quad (2.19)$$

Ainsi  $\gamma(\pi p/M^j)\delta[k]$  est une approximation de  $R_{j, n_j}^S[k]$  avec une marge de  $\pi/\omega_0 2M^j$ . Ce qui illustre la vitesse du processus de décorrélation, lorsque  $\omega_0 > \pi$ .

(b) Si  $\omega_0 \leq \pi$ , alors  $\gamma$  s'annule dans  $[\omega_0, \pi]$ . Si  $a(\kappa)$  est tel que  $\omega_0 < a(\kappa) \leq \pi$ , alors il existe  $\alpha > 0$  tel que  $\gamma$  s'annule sur  $]a(\kappa) - \alpha, a(\kappa) + \alpha[$  et il existe  $j_1$  tel que pour tout  $j \geq j_1$ ,  $\Delta_{j, p}^+ \subset ]a(\kappa) - \alpha, a(\kappa) + \alpha[$ . Par conséquent,  $R_{j, n}^S[k]$  s'annule aussi dès que  $j \geq j_1$ . Maintenant,



considérons que  $0 \leq a(\kappa) < \omega_0$ . Alors pour  $j$  supérieur ou égal à un certain  $j_2$ ,  $\Delta_{j,p}^+ \subset [0, \omega_0]$ . Ainsi, pour tout  $j \geq j_2$ ,  $\Delta_{j,p}^+ \cap [0, \omega_0] = \Delta_{j,p}^+$  et on obtient que l'Eq. (2.19) est vraie. Il est également trivial de vérifier que pour  $j$  supérieur ou égal à un certain  $j_3$ , l'Eq. (2.19) est vérifiée lorsque  $a(\kappa) = \omega_0$ . Pour résumé, si  $\omega_0 \leq \pi$ , il existe  $j_0$  tel que l'Eq. (2.19) soit également vraie pour tout  $j \geq j_0$ .

L'exemple de densité spectrale (Eq. (2.16)) considéré ci-dessus n'est pas anodin. En effet, considérons un spectre quelconque  $\gamma$  qui est continu et à support dans  $[-\pi, \pi]$ . Lorsque  $j$  est assez grand,  $\gamma$  peut être approchée par une fonction linéaire sur  $\Delta_{j,p}^+$ , compte tenu de la finesse de cet intervalle. On peut donc généraliser le résultat obtenu ci-dessus, et stipuler qu'au-delà d'un certain niveau de décomposition  $j_0$ ,  $\gamma(p\pi/M^j)\delta[k]$  est une approximation de  $R_{j,n}^S[k]$  avec une erreur inférieure ou égale à  $A/M^j$ . La constante  $A$  dépend de la pente de  $\gamma$  sur  $\Delta_{j,p}^+$ , pour  $j \geq j_0$ . D'autre part, on peut aussi remarquer que plus  $M$  est grand, plus on peut espérer une meilleure décorrélation. Ainsi, le cas  $M = 2$  est le plus défavorable au sens de l'Eq. (2.19).

## 2.3 Analyse asymptotique pour quelques familles de filtres

### 2.3.1 Convergence des fonctions d'autocorrélations

L'intérêt du Théorème 1 est théorique puisque les filtres associés à la  $M$ -TPOD de Shannon sont à support infini et ne sont donc pas très appropriés pour les applications pratiques. Dans le but d'obtenir des résultats du même ordre pour des filtres d'intérêt pratique, nous supposons que la  $M$ -TPOD est effectuée avec des filtres  $h_m^{[r]}$ ,  $m = 0, 1, \dots, M-1$ , qui dépendent d'un paramètre  $r$  tel que

$$\lim_{r \rightarrow \infty} H_m^{[r]} = H_m^S \quad (\text{p.p.}), \quad (2.20)$$

où  $H_m^{[r]}$  est la transformée de Fourier discrète de  $h_m^{[r]}$  et  $H_m^S$  est donné par l'Eq. (1.18). La paramètre  $r$  est appelé ordre des filtres de la  $M$ -TPOD. Lorsque  $r$  tend vers l'infini, les filtres de la  $M$ -TPOD de réponses impulsionnelles  $\{h_m^{[r]}\}_{m=0,1,\dots,M-1}$  convergent dans le sens spécifié par l'Eq. (2.20) vers les filtres  $\{h_m^S\}_{m=0,1,\dots,M-1}$  de la  $M$ -TPOD de Shannon. D'autre part, l'Eq. (2.20) peut être vue comme étant une propriété de régularité pour les raisons suivantes. Selon [Saito and Beylkin, 1993; Shen and Strang, 1996, 1998], les filtres de Daubechies satisfont à l'Eq. (2.20) pour  $M = 2$  lorsque  $r$  est le nombre de moments nuls de la fonction d'ondelette de Daubechies. D'après [Aldroubi, Unser, and Eden, 1992], les filtres de Battle-Lemarié satisfont également à l'Eq. (2.20) pour  $M = 2$  lorsque  $r$  est l'ordre de la fonction d'échelle de Battle-Lemarié. L'existence de telles familles de filtres pour  $M > 2$  n'est pas encore formellement prouvée. Cependant, il est raisonnable de penser que les filtres de  $M$ -TPOD de Daubechies et de Battle-Lemarié convergeront vers les filtres de la  $M$ -TPOD de Shannon dans le sens précisé ci-dessus. On a le résultat suivant, qui concerne ces familles particulières de filtres.

**Théorème 2** *Soit  $X$ , un processus centré, du second ordre, stationnaire au sens large, et continu en moyenne quadratique. On suppose que le spectre  $\gamma$  de  $X$  est un élément de  $L^\infty(\mathbb{R})$  et est à support dans  $[-\pi, \pi]$ . Considérons la  $M$ -TPOD de l'espace  $\mathbf{U}^S$  de PW avec des filtres  $h_m^{[r]}$ ,  $m = 0, 1, \dots, M-1$ , satisfaisant à l'Eq. (2.20).*

Pour tout entier naturel  $j$  et tout  $n = 0, 1, \dots, M^j - 1$ , soit  $R_{j,n}^{[r]}$  la fonction d'autocorrélation des coefficients de  $X$  par projection sur  $\mathbf{W}_{j,n}^{[r]}$ .

(i) On a

$$\lim_{r \rightarrow +\infty} R_{j,n}^{[r]}[k] = R_{j,n}^S[k], \quad (2.21)$$

uniformément en  $k \in \mathbb{Z}$  et  $n$ , où  $R_{j,n}^S$  est donné par l'Eq. (2.12).

(ii) Soit  $\mathcal{P}_\kappa = (\mathbf{U}^S, \{\mathbf{W}_{j,n_\kappa(j)}^{[r]}\}_{j \in \mathbb{N}})$ , un chemin de la  $M$ -TPOD de Shannon, avec  $\kappa$ , une séquence d'éléments à valeur dans  $\{0, 1, \dots, M-1\}$ . Si  $\gamma$  est continu au point  $a(\kappa)$ , alors

$$\lim_{j \rightarrow +\infty} \left( \lim_{r \rightarrow +\infty} R_{j,n_\kappa(j)}^{[r]}[k] \right) = \gamma(a(\kappa))\delta[k], \quad (2.22)$$

uniformément en  $k \in \mathbb{Z}$ , avec  $a(\kappa)$  donné par l'Eq. (2.10).

**Preuve :** La preuve détaillée est donnée dans [Atto and Pastor, 2008b]. Remarquons simplement que puisque la  $M$ -TPOD concerne l'espace  $\mathbf{U}^S$ , alors on a

$$\mathcal{F}W_n^S(\omega) = \left[ \prod_{\ell=1}^j H_{m_\ell}^S\left(\frac{\omega}{M^{j+1-\ell}}\right) \right] \mathcal{F}\Phi^S\left(\frac{\omega}{M^j}\right), \quad (2.23)$$

et

$$\mathcal{F}W_n^{[r]}(\omega) = \left[ \prod_{\ell=1}^j H_{m_\ell}^{[r]}\left(\frac{\omega}{M^{j+1-\ell}}\right) \right] \mathcal{F}\Phi^S\left(\frac{\omega}{M^j}\right), \quad (2.24)$$

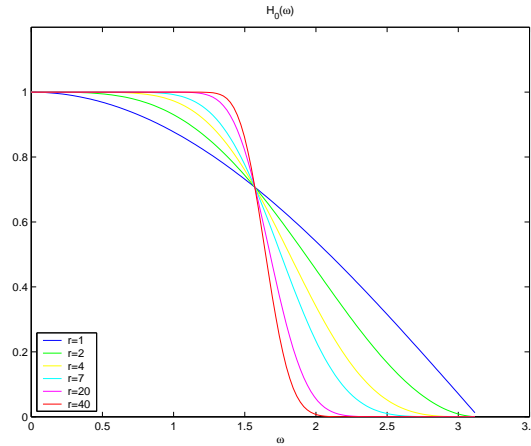
où  $n$  est donné par l'Eq. (1.15). Ces expressions résultent de l'Eq. (1.17) du lemme 1. À partir Eqs. (2.9), (2.20), (2.23) et (2.24), on déduit que pour tout entier naturel  $j$

$$\lim_{r \rightarrow +\infty} \mathcal{F}W_{j,n}^{[r]} = \mathcal{F}W_{j,n}^S \quad (\text{p.p.}), \quad (2.25)$$

uniformément en  $n$ . Ainsi,  $\mathcal{F}W_{j,n}^{[r]}$  converge presque partout vers  $\mathcal{F}W_{j,n}^S$  lorsque  $r \rightarrow +\infty$ , et cela explique en partie, les résultats énoncés par le théorème. ■

**Remarque 1** Considérons le chemin des approximations  $\mathcal{P}_\kappa = \mathcal{P}_{\kappa_0}$ ,  $\kappa_0$  étant la séquence nulle, de l'arbre de la  $M$ -TPOD. D'après l'Eq. (2.7), la décorrélation asymptotique est simplement obtenue lorsque le niveau de décomposition augmente. Cependant, l'Eq. (2.22) laisse supposer que l'ordre des filtres de la  $M$ -TPOD jouera un certain rôle dans la rapidité du processus de décorrélation. Les résultats expérimentaux donnés dans la section suivante (voir aussi [Atto, Pastor, and Isar, 2007]) confirment ce fait lorsque la TPOD est réalisée avec les filtres de Daubechies.

La Section 2.2.2 a montré qu'on pouvait espérer une convergence très rapide en utilisant les filtres de la  $M$ -TPOD de Shannon. Or, la convergence établie par le Théorème 2 est conditionnée par la convergence de certaines familles de filtres de la  $M$ -TPOD vers ceux de Shannon. Ainsi donc, plus cette dernière convergence sera rapide, mieux on peut espérer une décorrélation rapide. On peut vérifier que dans le cas de la famille de filtres de la TPOD de Daubechies, la vitesse de convergence est assez appréciable, comme le montre la figure 2.1. Ce qui permet de supposer qu'on devrait obtenir des résultats assez satisfaisants avec les filtres de la TPOD de Daubechies. La section suivante confirme cette hypothèse.



**Figure 2.1** — Réponses fréquentielles des filtres d'échelles de Daubechies d'ordres  $r = 1, 2, 4, 7, 20$ , et  $40$ . Cette figure illustre la vitesse de convergence des filtres de Daubechies vers les filtres de Shannon. Les filtres d'échelles sont obtenus à partir de la routine `dbaux` du wavelet toolbox de MatLab.

### 2.3.2 Résultats expérimentaux

Les résultats théoriques présentés ci-dessus sont de nature asymptotique. L'objectif de cette section est de montrer le niveau de décorrélation que l'on peut espérer atteindre en pratique, lorsque le niveau de décomposition et l'ordre des filtres sont finis. Considérons un processus aléatoire  $X(t)$  satisfaisant aux hypothèses du Théorème 2. Comme conséquence du Théorème 2 il existe un certain niveau de décomposition  $j$  et un ordre de filtre  $r$  tels que l'Eq. (2.22) soit satisfaite lorsque la décomposition est initialisée avec les échantillons du processus. L'Eq. (2.22) nous informe aussi que la valeur de la fonction d'autocorrélation en 0 tend vers la valeur du spectre du processus d'entrée en un certain point qui dépend du chemin choisi. Les résultats expérimentaux qui suivent ont pour but d'illustrer ces deux comportements.

Les résultats expérimentaux présentés dans cette section concernent la TPOD standard (2-TPOD), la décomposition se faisant avec les filtres de Daubechies. Nous considérons la TPOD pour les raisons suivantes. En premier lieu, on sait que les filtres de Daubechies, pour  $M = 2$ , convergent vers les filtres de Shannon lorsque le nombre  $r$  de moment nul de la fonction d'ondelettes de Daubechies tend vers l'infini, tandis que le même type de convergence n'est pas encore formellement établi pour  $M > 2$ . En plus, l'inégalité (3.9) suggère que : plus  $M$  est grand, plus rapide sera la tendance à la décorrélation. Ainsi, en prenant  $M = 2$ , on considère, non pas seulement le cas le plus standard, mais également le cas le moins favorable compte tenu de l'Eq. (3.9).

Les tests sont effectués sur les échantillons  $X(1), X(2), \dots, X(N)$  d'un processus  $X$  Auto-Régressif (AR) en entrée de la TPOD. Ces échantillons sont tels que  $X(\ell) = \mu X(\ell-1) + (1-\mu)Z(\ell)$ , avec  $0 < \mu < 1$ , et  $\ell = 1, 2, \dots, N$ , pour  $N = 2^{20}$ . Les variables aléatoires  $(Z(\ell))_\ell$  sont indépendantes et identiquement distribuées (iid), la distribution étant Gaussienne de moyenne nulle et de variance unité. Les échantillons  $(X(k))_k$  sont obtenus par filtrage du processus  $(Z(\ell))_\ell$  par un filtre discret AR d'ordre 1 dont la fonction de transfert est  $(1-\mu)/(1-\mu z^{-1})$ . Le spectre du processus  $X(t)$  est donc

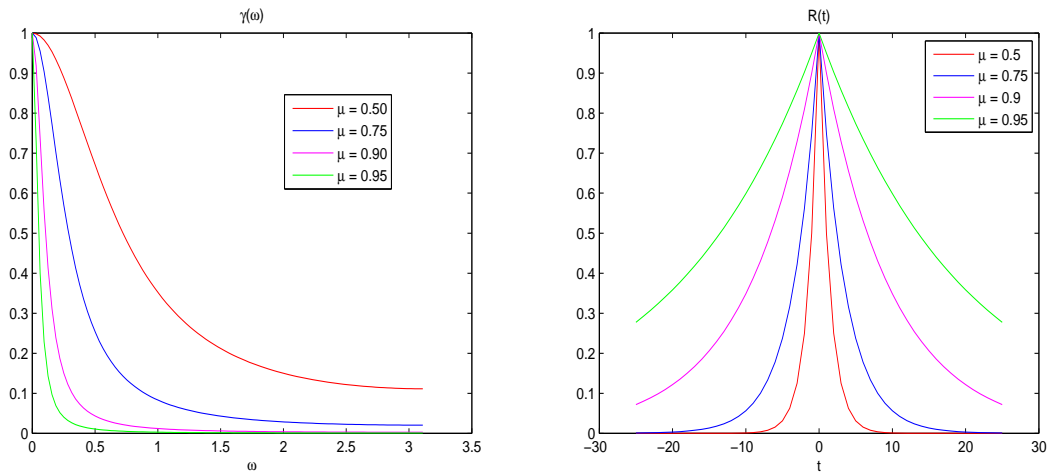
$$\gamma(\omega) = (1-\mu)^2 / |1 - \mu e^{-i\omega}|^2,$$

et sa fonction d'autocorrélation est

$$R[k] = \sigma^2 \frac{(1-\mu)^2}{1-\mu^2} \mu^{|k|},$$

où  $\sigma^2$  est la variance de  $Z(k)$ .

Le spectre de  $X(t)$ , ainsi que sa fonction d'autocorrélation normalisée, sont donnés en fonction de la valeur de  $\mu$  par la figure 2.2. Les valeurs considérées sont  $\mu = 0.5, 0.75, 0.9, 0.95$ . La



**Figure 2.2** — Spectres obtenues par filtrage AR(1) (gauche) et fonctions d'autocorrélation normalisées correspondantes(droite).

table 2.1 donne les valeurs  $\gamma(a(\kappa))$  pour 4 séquences tests qui sont les séquences présentées à l'Exemple 2. Les tests effectués montrent que le blanchiment (décorrélation) annoncé par le

**Tableau 2.1** — Valeurs de  $\gamma(a(\kappa))$  pour les quatre séquences tests.

Chemin	$\mu = 0.5$	$\mu = 0.75$	$\mu = 0.9$	$\mu = 0.95$
$\mathcal{P}_{\kappa_0}$	1	1	1	1
$\mathcal{P}_{\kappa_1}$	0.1111	0.0204	0.0028	0.0007
$\mathcal{P}_{\kappa_2}$	0.2052	0.0412	0.0057	0.0014
$\mathcal{P}_{\kappa_3}$	0.4798	0.1332	0.0201	0.0048

Théorème 2 est assez sensiblement atteint par la TPOD pour des valeurs raisonnables du niveau de décomposition et l'ordre des filtres d'analyse. Ces tests sont obtenus en considérant 100 réalisations du processus aléatoire décrit ci-dessus. La décomposition est faite avec les filtres de Daubechies et les résultats donnés ci-après sont les valeurs moyennes des tests effectués sur les 100 réalisations.

Nous commençons d'abord par les résultats illustrant l'influence du niveau de décomposition et l'ordre des filtres de la TPOD sur le processus de décorrélation. Ensuite, nous donnerons les résultats concernant les valeurs des fonctions d'autocorrélations en 0 (étude de la convergence de ces valeurs vers la variance du processus limite). Nous présentons les résultats obtenus avec  $\mu = 0.5$  et  $\mu = 0.9$ .

### Influence du niveau de décomposition et de l'ordre des filtres de la TPOD sur le processus de décorrélation

La TPOD est réalisée avec les filtres de Daubechies d'ordres  $1, 2, \dots, 20$ . Les fonctions d'autocorrélations moyennes des coefficients  $c_{j,n(\kappa_q)}$  pour  $q = 0, 1, 2, 3$ , (Exemple 2, Section 2.2.1) sont calculées sur un nombre de 1500 coefficients par paquet  $\mathbf{W}_{j,n(\kappa_q)}$ , ce qui est possible pour chaque niveau de décomposition puisque nous avons considéré  $2^{20}$  échantillons en entrée de la TPOD.

Il en résulte que pour les processus AR considérés, les coefficients de la TPOD peuvent être considérés comme assez sensiblement décorrélés lorsque le niveau de décomposition  $j$  est égal à 6 et l'ordre des filtres d'analyse  $r$  est égal à 7. La figure 2.4 montre les fonctions d'autocorrélations normalisées (afin de pouvoir apprécier le gain en décorrélation). Les graphes de la figure 2.4 montrent l'influence de  $j$  et  $r$  sur le processus de décorrélation.

L'importance de  $j$  se mesure par le fait que les coefficients obtenus au niveau de décomposition  $j = 6$  sont moins corrélés que ceux obtenus au niveau  $j = 3$  (comparer la première colonne de la figure 2.4 à la deuxième colonne de la même figure).

L'importance de  $r$  se juge par le fait que les coefficients obtenus par les filtres de Daubechies avec un ordre  $r = 7$  sont moins corrélés que ceux obtenus par les filtres de Daubechies avec un ordre  $r = 1$  (comparer la deuxième colonne de la figure 2.4 à la troisième colonne de la même figure).

### 2.3.3 Convergence des Valeurs en 0 des fonctions d'autocorrélations vers la variance du processus limite

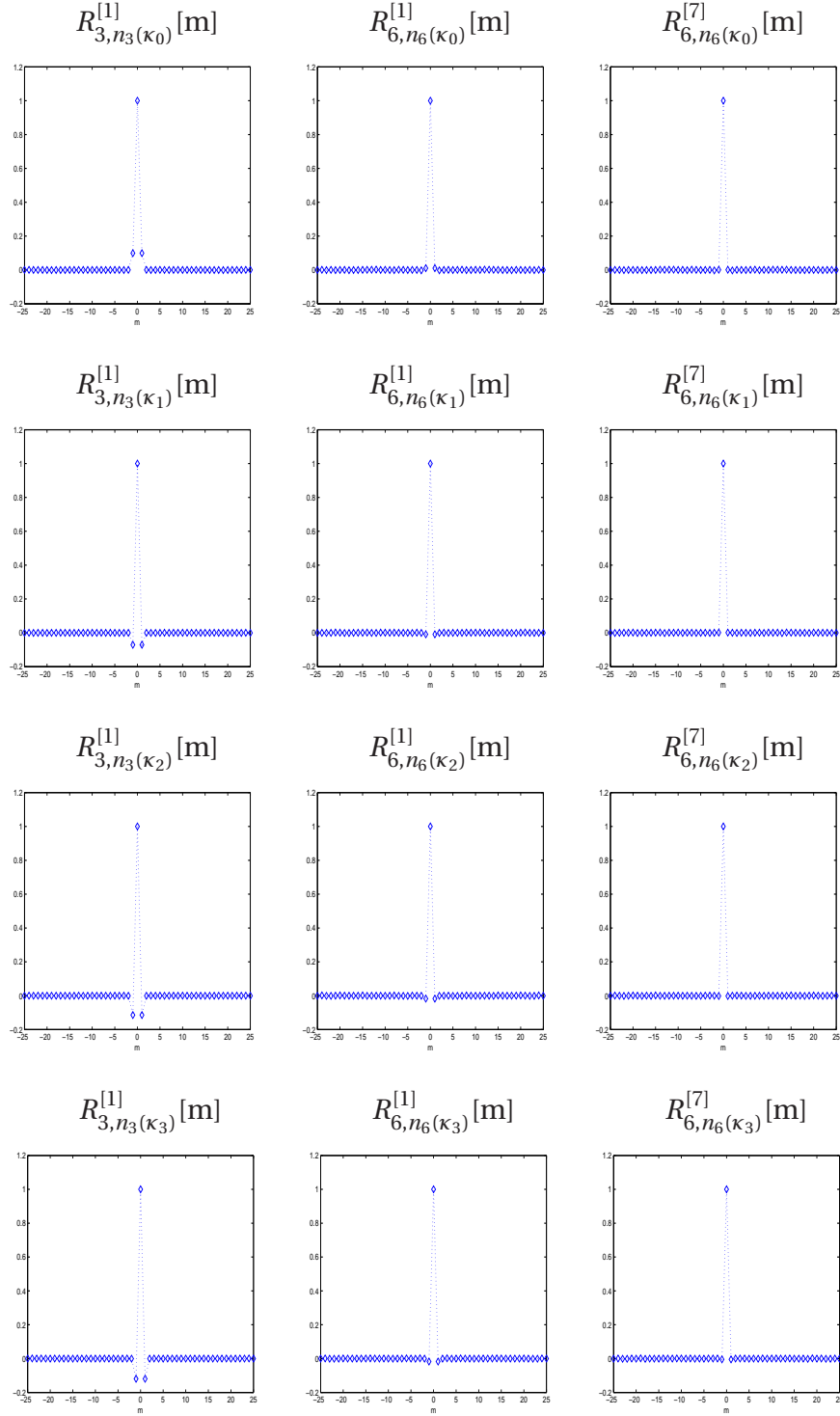
Considérons une séquence  $(R_{j,n}^{[r]})_j$  associée à un chemin donné dans l'arbre de la TPOD. On rappelle que pour tout entier naturel  $j$ ,  $R_{j,n}^{[r]}$  est la fonction d'autocorrélation des coefficients obtenus par projection de  $X$  sur le paquet  $\mathbf{W}_{j,n}^{[r]}$ . D'après le Théorème 2, la valeur  $R_{j,n}^{[r]}[k]$  sera proche de  $\gamma(a(\kappa))\delta[k]$  si  $j$  et  $r$  sont suffisamment grands. Nous rappelons que  $\gamma(a(\kappa))$  est la valeur du spectre du processus aléatoire  $X$  au point  $a(\kappa)$  où  $a(\kappa)$  est donné par l'Eq. (2.10) et  $\kappa$  est la séquence binaire associée à  $(\mathbf{W}_{j,n}^{[r]})_j$  et spécifiant un chemin donné dans l'arbre de la TPOD.

Nous allons illustrer ce résultat en considérant les valeurs en 0 de la séquence des fonctions d'autocorrélations associées aux séquences tests  $\kappa_0, \kappa_1, \kappa_2$ , et  $\kappa_3$  introduites dans l'exemple 2. Nous nous intéressons aux filtres de la TPOD d'ordre  $r = 7$  puisque cet ordre peut être qualifié de suffisamment grand compte tenu des résultats expérimentaux donnés ci-dessus.

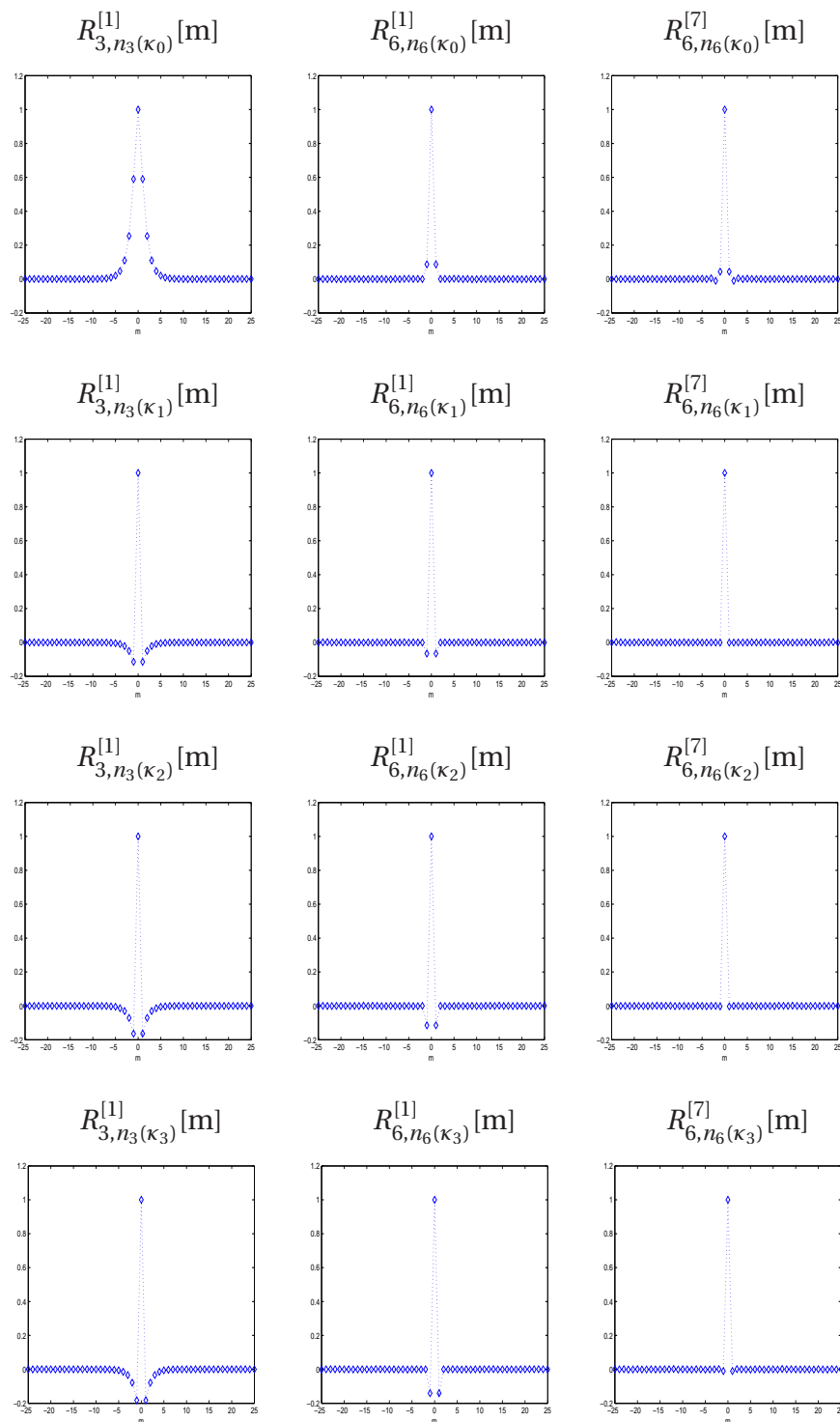
Les tests concernent le processus AR d'ordre 1 décrit dans la section précédente et dont les spectres sont donnés en fonction de la valeur de  $\mu$  par la figure 2.2. On observe que les termes de la séquence  $(R_{j,n}^{[7]}[0])_j$  deviennent assez rapidement proche de la limite  $\gamma(a(\kappa))$  lorsque  $j$  augmente. Cela est illustré par la figure 2.5. Les graphes de la figure 2.5 représentent les valeurs moyennes  $R_{j,n_{\kappa_q(j)}}^{[7]}[0]$  pour  $j = 0, 1, 2, \dots, 9$  et pour les 4 séquences  $\kappa_q$ ,  $q = 0, 1, 2, 3$ . La valeur limite  $\gamma(a(\kappa_q))$  est tracée en pointillés sur chaque graphe. On observe que, lorsque  $j$  augmente jusqu'à atteindre 6, la valeur en 0 de la fonction d'autocorrélation des coefficients de la TPOD est assez proche de la valeur  $\gamma(a(\kappa_q))$ , pour  $q = 0, 1, 2, 3$ . Notons enfin que l'ap-

précision de la précision de ces résultats est donnée dans [Atto, Pastor, and Isar, 2007] pour les valeurs moyennes  $R_{j, n_{\kappa_q}(j)}^{[7]}[0]$  lorsque  $j = 6, 7, 8, 9$  par des intervalles de confiance à 95% calculés à l'aide du t-test standard.

Pour conclure cette section, on peut noter que le spectre du processus AR a une pente très forte au voisinage de 0 lorsque  $0.9 \leq \mu < 1$ . Or, d'après l'analyse faite à la Section 2.2.2, on sait que la raideur de cette pente est de nature à ralentir considérablement le processus de blanchiment. En effet, plus le spectre est plat, plus rapidement on obtient des séquences de coefficients décorrelés. Ainsi, compte tenu du fait que les résultats présentés sont assez satisfaisants (pour les coefficients de la séquence  $\kappa_0$  concernée par la valeur en 0 du spectre du processus AR de paramètre  $\mu = 0.9$ ), il est légitime de penser pour un processus quelconque dont le spectre n'est pas plus singulier que celui du processus AR de paramètre  $\mu = 0.9$ , on atteint raisonnablement la décorrélation dans la plupart des paquets d'ondelettes, lorsqu'on se situe au sixième niveau de décomposition (ou au delà) et lorsque les filtres de la TPOD sont d'ordre supérieur ou égal à 7.

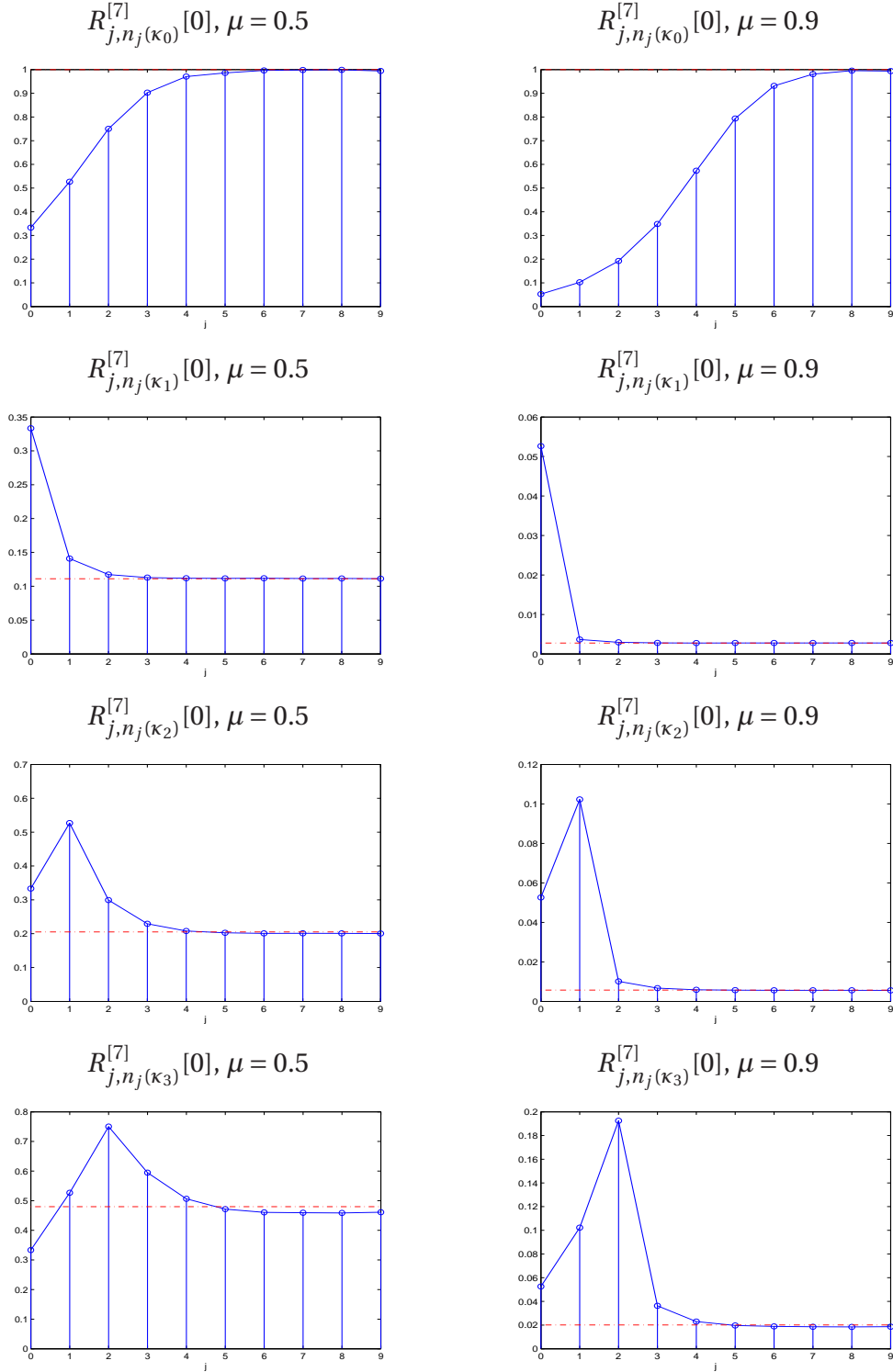


**Figure 2.3** — Fonctions d'autocorrélation normalisées des coefficients des paquets d'ondelettes obtenus pour les 4 séquences de l'exemple 1, aux niveaux de décomposition 3 et 6. Les décompositions concernent un processus aléatoire obtenu par filtrage d'un bruit blanc Gaussien avec un filtre AR(1) de paramètre  $\mu = 0.5$  et dont le spectre est donné à la figure 2.2. Elles sont effectuées avec les filtres de Daubechies d'ordres 1 et 7.



**Figure 2.4** — Fonctions d'autocorrélation normalisées des coefficients des paquets d'ondelettes obtenus pour les 4 séquences de l'exemple 1, aux niveaux de décomposition 3 et 6. Les décompositions concernent un processus aléatoire obtenu par filtrage d'un bruit blanc Gaussien avec un filtre AR(1) de paramètre  $\mu = 0.9$  et dont le spectre est donné à la figure 2.2. Elles sont effectuées avec les filtres de Daubechies d'ordres 1 et 7.





**Figure 2.5** — Convergence de  $R_{j,n_j(\kappa_q)}^{[7]}[0]$  vers  $\gamma(a(\kappa_q))$  avec  $j$  pour les quatre séquences tests. La décomposition est effectuée avec les filtres de Daubechies d'ordre 7 et concerne des processus aléatoires obtenus par filtrage de bruit Gaussien blanc avec des filtres AR(1) de paramètres  $\mu = 0.5$  et  $\mu = 0.9$ . Les spectres de ces processus sont données à la figure 2.2.

# Analyse asymptotique des distributions de la $M$ -TPOD de processus stationnaires au sens strict

## 3.1 Rappels - État de l'art

Dans cette section, nous considérons un processus aléatoire  $X$  admettant des cumulants et polyspectres finis. Notons

$$\text{cum}(t, s_1, s_2, \dots, s_N) = \text{cum}\{X(t), X(s_1), X(s_2), \dots, X(s_N)\}, \quad (3.1)$$

le cumulant d'ordre  $N + 1$  de  $X$ . Ce cumulant est supposé appartenir à  $L^2(\mathbb{R}^{N+1})$  et être fini pour tout entier naturel  $N$  (voir [Touati and Pesquet, 2002, Proposition 1] pour une discussion au sujet de l'existence de ce cumulant). Le cumulant d'ordre  $N + 1$  du processus  $c_{j,n}$  a la forme intégrale donnée par (voir [Touati and Pesquet, 2002, Proposition 1]) :

$$\begin{aligned} \text{cum}_{j,n}[k, \ell_1, \ell_2, \dots, \ell_N] &= \text{cum}\{c_{j,n}[k]c_{j,n}[\ell_1]c_{j,n}[\ell_2]\cdots c_{j,n}[\ell_N]\}, \\ &= \int_{\mathbb{R}^{N+1}} dt ds_1 ds_2 \cdots ds_N \text{cum}(t, s_1, s_2, \dots, s_N) \\ &\quad W_{j,n,k}(t) W_{j,n,\ell_1}(s_1) W_{j,n,\ell_2}(s_2) \cdots W_{j,n,\ell_N}(s_N). \end{aligned} \quad (3.2)$$

En posant  $s_i = t + t_i$  pour  $i = 1, 2, \dots, N$  dans l'Eq. (3.2) et en supposant que  $X$  est strictement stationnaire de sorte que  $\text{cum}(t, t + t_1, t + t_2, \dots, t + t_N) = \text{cum}(t_1, t_2, \dots, t_N)$ , alors

$$\begin{aligned} \text{cum}_{j,n}[k, \ell_1, \ell_2, \dots, \ell_N] &= \int_{\mathbb{R}^{N+1}} dt dt_1 dt_2 \cdots dt_N \text{cum}(t_1, t_2, \dots, t_N) W_{j,n,k}(t) \\ &\quad W_{j,n,\ell_1}(t + t_1) W_{j,n,\ell_2}(t + t_2) \cdots W_{j,n,\ell_N}(t + t_N). \end{aligned} \quad (3.3)$$

Dans ce qui suit,  $X$  est supposé avoir un polyspectre  $\gamma_N = \gamma_N(\omega_1, \omega_2, \dots, \omega_N)$  pour tout entier naturel  $N$  et tout  $(\omega_1, \omega_2, \dots, \omega_N) \in \mathbb{R}^N$ . Le polyspectre  $\gamma_N$  est la transformée de Fourier du cumulant  $\text{cum}(t_1, t_2, \dots, t_N)$ . Lorsque  $N = 1$ ,  $\gamma_1$  est le spectre de  $X$  et est simplement noté  $\gamma$  comme c'est le cas dans le chapitre 2. Nous supposons dans toute la suite de ce chapitre que  $\gamma_N$  appartient à  $L^\infty(\mathbb{R}^N)$  et que

$$\int_{\mathbb{R}^N} d\omega_1 d\omega_2 \cdots d\omega_N \mathcal{F}W_n(-\omega_1 - \omega_2 - \cdots - \omega_N) \mathcal{F}W_n(\omega_1) \cdots \mathcal{F}W_n(\omega_N)$$

existe et est fini pour tout entier naturel  $N$ . Une condition suffisante pour l'existence de cette intégrale concerne la régularité de la fonction  $\Phi$  utilisée pour générer les fonctions  $W_n$  des paquets d'ondelettes. Cette condition est l'existence de deux constantes réelles positives  $\varepsilon$  et  $C$  tel que

$$|\mathcal{F}\Phi(\omega)| \leq \frac{C}{1 + |\omega|^{1+\varepsilon}}, \quad \omega \in \mathbb{R}. \quad (3.4)$$

L'Eq. (3.3) peut alors s'écrire (voir [Atto and Pastor, 2008b])

$$\begin{aligned} \text{cum}_{j,n}[k, \ell_1, \ell_2, \dots, \ell_N] &= \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \dots d\omega_N \\ &\quad \gamma_N(-\omega_1, -\omega_2, \dots, -\omega_N) \\ &\quad \mathcal{F}W_{j,n,k}(-\omega_1 - \omega_2 - \dots - \omega_N) \\ &\quad \mathcal{F}W_{j,n,\ell_1}(\omega_1) \mathcal{F}W_{j,n,\ell_2}(\omega_2) \dots \mathcal{F}W_{j,n,\ell_N}(\omega_N). \end{aligned} \quad (3.5)$$

En outre, de la première égalité de l'Eq. (1.10), à savoir  $W_{j,n,k}(t) = W_{j,n}(t - M^j k)$ , on a  $\mathcal{F}W_{j,n,k}(\omega) = e^{-iM^j k \omega} \mathcal{F}W_{j,n}(\omega)$ . Ainsi, en posant  $\ell_i = k + k_i$  pour  $i = 1, 2, \dots, N$  dans l'Eq. (3.5), on obtient que le processus aléatoire discret  $c_{j,n}$  est strictement stationnaire de cumulants  $\text{cum}_{j,n}[k, k + k_1, k + k_2, \dots, k + k_N] = \text{cum}_{j,n}[k_1, k_2, \dots, k_N]$  avec

$$\begin{aligned} \text{cum}_{j,n}[k_1, k_2, \dots, k_N] &= \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \dots d\omega_N \\ &\quad e^{-iM^j(k_1\omega_1 + k_2\omega_2 + \dots + k_N\omega_N)} \\ &\quad \gamma_N(-\omega_1, -\omega_2, \dots, -\omega_N) \\ &\quad \mathcal{F}W_{j,n}(-\omega_1 - \omega_2 - \dots - \omega_N) \\ &\quad \mathcal{F}W_{j,n}(\omega_1) \mathcal{F}W_{j,n}(\omega_2) \dots \mathcal{F}W_{j,n}(\omega_N). \end{aligned} \quad (3.6)$$

La nature asymptotique de la distribution du processus aléatoire obtenu au nœud  $(j, n)$  sera analysée à travers le comportement des cumulants  $\text{cum}_{j,n}[k_1, k_2, \dots, k_N]$ . En particulier, il sera prouvé que ces cumulants tendent vers zéro sous certaines conditions concernant les filtres d'analyse de la  $M$ -TPOD.

Compte tenu de l'Eq. (2.9), l'Eq. (3.6) peut encore s'écrire, (après un changement de variables) :

$$\begin{aligned} \text{cum}_{j,n}[k_1, k_2, \dots, k_N] &= \frac{M^{-j(N-1)/2}}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \dots d\omega_N \\ &\quad e^{-i(k_1\omega_1 + k_2\omega_2 + \dots + k_N\omega_N)} \\ &\quad \gamma_N\left(-\frac{\omega_1}{M^j}, -\frac{\omega_2}{M^j}, \dots, -\frac{\omega_N}{M^j}\right) \\ &\quad \mathcal{F}W_n(-\omega_1 - \omega_2 - \dots - \omega_N) \\ &\quad \mathcal{F}W_n(\omega_1) \mathcal{F}W_n(\omega_2) \dots \mathcal{F}W_n(\omega_N). \end{aligned} \quad (3.7)$$

Si le paramètre de décalage  $n$  est supposé indépendant de  $j$ , on obtient par application du théorème de la convergence dominée de Lebesgue que, pour tout entier naturel  $N > 1$ ,  $\text{cum}_{j,n}[k_1, k_2, \dots, k_N]$  tend vers 0 uniformément en  $k_1, k_2, \dots, k_N$  lorsque  $j$  tend vers  $\infty$ . Ce résultat est une conséquence de [Leporini and Pesquet, 1999, Proposition 11].

Cependant, on a vu (conséquence de l'Eq. (1.16)) que  $n$  ne pouvait être supposé indépendant de  $j$  que pour le chemin des approximations. Pour tous les autres chemins,  $n$  dépend

explicitement de  $j$  et ne pouvait être majoré par une constante indépendante de  $j$ . Par conséquent, lorsque le paramètre de décalage  $n$  dépend explicitement de  $j$ , le théorème de la convergence dominée de Lebesgue ne peut pas être appliqué à l'Eq. (3.7) pour prouver la convergence vers 0 des cumulants en fonction de  $j$ . Cependant, le Théorème 3 suivant montre cette convergence vers 0 des cumulants  $\text{cum}_{j,n}[k_1, k_2, \dots, k_N]$  pour  $N > 1$  lorsque la  $M$ -TPOD concerne les filtres de Shannon. dans l'hypothèse où la  $M$ -TPOD concerne les filtres de Shannon. Le Théorème 4 fournit un résultat du même ordre que le Théorème 3 dans l'hypothèse où la  $M$ -TPOD utilise les familles de filtres décrits au chapitre précédent, celles dont les transformées de Fourier de leurs filtres convergent presque partout vers les transformées de Fourier des filtres de Shannon.

## 3.2 Distributions asymptotiques des coefficients de la $M$ -TPOD

Nous avons le résultat suivant :

**Théorème 3** *Soit  $X$ , un processus centré, du second ordre, stationnaire au sens strict, et continu en moyenne quadratique. On suppose que le polyspectre  $\gamma_N$  de  $X$  est un élément de  $L^\infty(\mathbb{R}^N)$  pour tout  $N \geq 1$  et que le spectre  $\gamma$  est à support dans  $[-\pi, \pi]$ . Pour tout entier naturel  $N > 1$ , le cumulants d'ordre  $N+1$ ,  $\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]$ , du processus aléatoire discret obtenu au nœud  $(j, n)$  par la  $M$ -TPOD de Shannon de  $X$  tend vers zéro uniformément en  $n, k_1, k_2, \dots, k_N$ , lorsque  $j$  tend vers l'infini.*

La preuve de ce résultat est donnée dans [Atto and Pastor, 2008b]. Nous la rappelons, compte tenu de sa simplicité.

**Preuve :** Lorsque la décomposition concerne les fonctions d'ondelettes  $W_{j,n}^S$  de Shannon, on obtient des Eqs. (1.21) et (3.6) que le cumulants  $\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]$  du processus aléatoire discret obtenu au nœud  $(j, n)$  par la  $M$ -TPOD de Shannon de  $X$  satisfait l'inégalité suivante

$$|\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]| \leq \frac{M^{j(N+1)/2}}{(2\pi)^N} \cdot \|\gamma_N\|_\infty \int_{\Delta_{j,p}^N} d\omega_1 d\omega_2 \dots d\omega_N, \quad (3.8)$$

où  $\Delta_{j,p}^N = \underbrace{\Delta_{j,p} \times \Delta_{j,p} \times \dots \times \Delta_{j,p}}_{N \text{ fois}}$  et  $p = G(n)$ .

Mais on a  $\int_{\Delta_{j,p}} d\omega = 2\pi / M^j$  compte tenu de l'Eq. (1.22). On obtient donc

$$|\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]| \leq \|\gamma_N\|_\infty M^{-j(N-1)/2}. \quad (3.9)$$

Pour tout entier naturel  $N > 1$ , le majorant  $\|\gamma_N\|_\infty M^{-j(N-1)/2}$  de  $|\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]|$  ne dépend pas de  $n, k_1, \dots, k_N$  et tend vers 0 lorsque  $j$  tend vers l'infini, d'où le résultat. ■

Considérons alors les filtres de la  $M$ -TPOD introduits dans la Section 2.2.1 et satisfaisant donc à l'Eq. (2.20). On a le résultat suivant.

**Théorème 4** Soit  $X$ , un processus centré, du second ordre, stationnaire au sens strict, et continu en moyenne quadratique. On suppose que le polyspectre  $\gamma_N$  de  $X$  est un élément de  $L^\infty(\mathbb{R}^N)$  pour tout  $N \geq 1$  et que le spectre  $\gamma$  est à support dans  $[-\pi, \pi]$ .

Pour tout entier naturel  $j$  donné et tout  $n \in \{0, 1, \dots, M^j - 1\}$ , soit  $\text{cum}_{j,n}^{[r]}$  le cumulante d'ordre  $N + 1$  des coefficients de la  $M$ -TPOD de  $X$  par projection sur  $\mathbf{W}_{j,n}^{[r]}$ .

(i) On a

$$\lim_{r \rightarrow +\infty} \text{cum}_{j,n}^{[r]}[k_1, k_2, \dots, k_N] = \text{cum}_{j,n}^S[k_1, k_2, \dots, k_N], \quad (3.10)$$

uniformément en  $n, k_1, k_2, \dots, k_N$ .

(ii) Soit  $\kappa = (m_k)_{k \in \mathbb{N}}$  une séquence d'éléments à valeurs dans  $\{0, 1, \dots, M - 1\}$  et  $\mathcal{P}_\kappa = (\mathbf{U}^S, \{\mathbf{W}_{j,n_\kappa(j)}^{[r]}\}_{j \in \mathbb{N}})$  la séquence de paquets d'ondelettes associée à  $\kappa$ . Si  $N > 1$ , on a

$$\lim_{j \rightarrow +\infty} \left( \lim_{r \rightarrow +\infty} \text{cum}_{j,n_\kappa(j)}^{[r]}[k_1, k_2, \dots, k_N] \right) = 0, \quad (3.11)$$

uniformément en  $k_1, k_2, \dots, k_N$ .

**Preuve :** [Atto and Pastor, 2008b]. ■

**Remarque 2** Si  $n = n_\kappa(j)$  est constant avec  $j$ , donc si  $\kappa = \kappa_0$  est en fait la séquence nulle associée au chemin des approximations, le comportement asymptotique du cumulante  $\text{cum}_{j,n}^{[r]}[k_1, k_2, \dots, k_N]$  lorsque  $j$  tend vers l'infini est une conséquence de l'Eq. (3.7). Cependant, de manière analogue à la remarque 1, l'Eq. (3.11) suggère que l'ordre  $r$  jouera un rôle dans la convergence vers 0 des cumulants. Cette hypothèse est vérifiée pour les résultats expérimentaux qui suivent car les tests effectués montrent que pour le chemin  $\kappa_0$ , la convergence s'accélère lorsque l'ordre des filtres augmente.

Les résultats suivants décrivent la nature asymptotique des distributions des processus aléatoires  $c_{j,n}^S$  et  $c_{j,n}^{[r]}$ .

**Théorème 5** Avec les notations et hypothèses du Théorème 4, et si  $\gamma$  est continue au point  $a(\kappa)$ , alors la séquence  $(c_{j,n_\kappa(j)}^S)_j$  converge en distribution vers un processus blanc et Gaussien de variance  $\gamma(a(\kappa))$  lorsque  $j$  tendent vers l'infini. La convergence a lieu dans le sens suivant. Pour tout  $x \in \mathbb{R}^N$  et tout  $\epsilon > 0$ , il existe  $j_0 = j_0(x, \epsilon) > 0$  tel que, pour tout  $j \geq j_0$ , la valeur absolue de la différence entre la valeur en un point  $x$  de la distribution de probabilité du vecteur aléatoire

$$(c_{j,n_\kappa(j)}^S[k_1], c_{j,n_\kappa(j)}^S[k_2], \dots, c_{j,n_\kappa(j)}^S[k_N])$$

et la valeur au même point  $x$  de la distribution Gaussienne centrée  $N$ -dimensionnelle de matrice de covariance  $\gamma(a(\kappa))\mathbf{I}_N$ , notée  $\mathcal{N}(0, \gamma(a(\kappa))\mathbf{I}_N)$ , est inférieure à  $\epsilon$ .

**Preuve :** Ce résultat est une conséquence du Théorème 1 et du Théorème 3. ■

**Théorème 6** Avec les notations et hypothèses du Théorème 4, et si  $\gamma$  est continue au point  $a(\kappa)$  alors la séquence  $\left(c_{j,n_\kappa(j)}^{[r]}\right)_{r,j}$  converge en distribution vers un processus blanc et Gaussien de variance  $\gamma(a(\kappa))$  lorsque  $j$  et  $r$  tendent vers l'infini : Pour tout  $x \in \mathbb{R}^N$  et tout  $\epsilon > 0$ , il existe  $j_0 = j_0(x, \epsilon) > 0$  et  $r_0 = r_0(x, j_0, \epsilon)$  tels que, pour tout  $j \geq j_0$  et tout  $r \geq r_0$ , la valeur absolue de la différence entre la valeur en un point  $x$  de la distribution de probabilité du vecteur aléatoire

$$(c_{j,n_\kappa(j)}^{[r]}[k_1], c_{j,n_\kappa(j)}^{[r]}[k_2], \dots, c_{j,n_\kappa(j)}^{[r]}[k_N])$$

et la valeur au même point  $x$  de la distribution  $\mathcal{N}(0, \gamma(a(\kappa))I_N)$  est inférieure à  $\epsilon$ .

**Preuve :** Ce résultat est une conséquence de la condition (ii) du Théorème 2 et de la condition (ii) du Théorème 4. ■

La section suivante étudie la portée des résultats énoncés ci-dessus, pour des applications pratiques, lorsque le niveau de décomposition et l'ordre des filtres d'analyse sont finis.

### 3.3 Résultats expérimentaux

Les résultats expérimentaux présentés ici concernent la TPOD standard, lorsque les filtres de Daubechies sont utilisés pour la décomposition. Ces résultats ont pour objectif de vérifier la tendance à la *normalité* décrite par le Théorème 4. L'inégalité (3.9) suggère que plus  $M$  est grand, plus on peut espérer une vitesse de convergence rapide. Ainsi, considérer la TPOD ( $M = 2$ ) revient donc à considérer le cas le moins favorable.

On considère un processus aléatoire Hilbertien  $X(t)$  en entrée de la TPOD. Le spectre de ce processus est noté  $\gamma$ , est supposé appartenir à  $L^\infty(\mathbb{R})$ , et est à support dans  $[-\pi, \pi]$ .

Les tests effectués concernent le cas où  $X(t)$  est un processus Gaussien Généralisé (GG). Cela signifie que pour tout  $t \in \mathbb{R}$ ,  $X(t)$  suit une distribution GG d'échelle  $\alpha$ , de forme  $\beta$  et de moyenne nulle. Pour tout  $t \in \mathbb{R}$ , la densité de probabilité (loi) de  $X(t)$  est  $f_{\alpha,\beta}$ , définie pour tout  $x$  réel par :

$$f_{\alpha,\beta}(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta}, \quad (3.12)$$

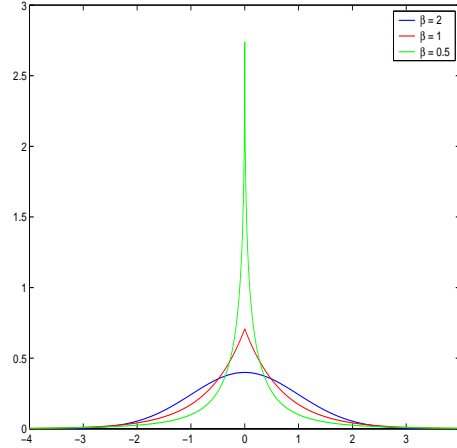
où  $\Gamma$  est la fonction Gamma standard. La variance de la distribution GG est

$$\sigma = \alpha \sqrt{\Gamma(3/\beta)/\Gamma(1/\beta)}.$$

Nous posons

$$\alpha = \sqrt{\Gamma(1/\beta)/\Gamma(3/\beta)}, \quad (3.13)$$

de sorte que  $\sigma = 1$  dans tous les tests réalisés. Notons que lorsque le paramètre de forme  $\beta$  de la distribution GG est égal à 2, cette distribution est Gaussienne (remplacer  $\beta$  par 2 dans l'Eq. (3.12)). Lorsque  $\beta$  décroît (de 2 à 0), la densité de la distribution GG devient de plus en plus singulière (pointue) en 0, comme on peut le voir sur la figure 3.1. lorsque  $\beta = 1$ , la GG est une distribution Laplacienne.



**Figure 3.1** — Distributions GGs de variance unité et de formes  $\beta = 2, 1, 0.5$ .

Pour réaliser les tests expérimentaux, nous considérons 100 copies indépendantes du vecteur formé par  $N$  échantillons  $X(1), X(2), \dots, X(N)$  avec  $N = 2^{20}$ . Chaque copie est une entrée de la TPOD. Nous considérons également les chemins de la TPOD associés aux 4 séquences  $\kappa_q = (\delta[q - \ell])_{\ell \in \mathbb{N}}$ , pour  $q = 0, 1, 2$  et 3 (voir exemple 2). On rappelle que pour la séquence nulle  $\kappa_0$ , on a  $n_{\kappa_0}(\ell) = 0$  pour tout entier naturel  $\ell$ , et que pour les autres séquences  $\kappa_q$ ,  $q = 1, 2, 3$ , on a :

$$n_{\kappa_q}(\ell) = \begin{cases} 0 & \text{for } \ell = 1, 2, \dots, q-1, \\ 2^{\ell-q} & \text{for } \ell = q, q+1, \dots. \end{cases}$$

et que,  $p_{\kappa_0}(\ell) = G(n_{\kappa_0}(\ell)) = 0$ ,

$$p_{\kappa_q}(\ell) = \begin{cases} 0 & \text{for } \ell = 1, 2, \dots, q-1, \\ 2^{\ell-q+1} - 1 & \text{for } \ell = q, q+1, \dots, \end{cases}$$

Ce qui conduit à  $a(\kappa_0) = 0$  et  $a(\kappa_q) = \pi/2^{q-1}$  pour  $q = 1, 2$  et 3.

Le test de Kolmogorov-Smirnov (test KS en abrégé) est un test d'hypothèses basé sur les fonctions de répartitions empiriques. Ce test sert à déterminer si un échantillon donné suit une loi précise ou non. Le test KS est utilisé dans cette section pour décider si la séquence  $(c_{j, n_{\kappa}(j)}^{[r]}[k] / \sqrt{\gamma(a(\kappa))})_k$  suit la loi normale  $\mathcal{N}(0, 1)$  (hypothèse nulle) ou non (hypothèse alternative), pour chacune des séquences  $\kappa$  considérée ci-dessus.

En premier, nous traitons le cas où les échantillons  $X(1), X(2), \dots, X(N)$  du processus GG  $X(t)$  sont non-corrélés. Dans ce cas, et puisque la distribution GG est telle que  $\sigma = 1$ , nous avons  $\gamma(a(\kappa)) = 1$  pour chacun des 4 chemins  $(\mathcal{P}_{\kappa_q})_{q=0,1,2,3}$  considérés. Les tests montrent qu'en augmentant le niveau de résolution  $j$  jusqu'à 6 et en utilisant des filtres de Daubechies d'ordre 7, la convergence vers la distribution normale  $\mathcal{N}(0, 1)$  est significative pour  $0.5 < \beta \leq 2$ . La table 3.1 présente les résultats obtenus pour  $\beta = 1.5, 1, 0.5$ .

Nous traitons maintenant le cas où les échantillons  $X(1), X(2), \dots, X(N)$  du processus GG  $X(t)$  sont corrélées. Ces échantillons sont synthétisés par filtrage d'une séquence de variables aléatoires iid GG à travers un filtre AR d'ordre 1, de sorte que le spectre de  $X(t)$  est

$$\gamma(\omega) = (1 - \mu)^2 / |1 - \mu e^{-i\omega}|^2,$$

**Tableau 3.1** — Taux d'acceptation par le test KS de la distribution normale  $\mathcal{N}(0, 1)$  pour les coefficients de la TPOD obtenus aux niveaux de décompositions  $j = 3, 6$  et pour différents chemins. La TPOD est réalisée avec les filtres de Daubechies d'ordres 1 et 7. Le processus d'entrée est GG. Les paramètres de la distribution GG sont :  $\alpha$  donné par l'Eq. (3.13) et  $\beta = 1.5, 1, 0.5$ . Les échantillons du processus GG sont décorrélés.

Processus GG blanc									
Chemin	GG, avec $\beta = 1.5$			GG, avec $\beta = 1$			GG, avec $\beta = 0.5$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\kappa_0}$	51%	94%	96%	0%	97%	97%	0%	44%	74%
$\mathcal{P}_{\kappa_1}$	53%	97%	98%	0%	96%	93%	0%	44%	71%
$\mathcal{P}_{\kappa_2}$	47%	94%	97%	0%	94%	96%	0%	31%	39%
$\mathcal{P}_{\kappa_3}$	56%	94%	94%	0%	92%	94%	0%	37%	38%

avec  $0 < \mu < 1$  (voir figure 2.2). Si  $\alpha$  et  $\beta$  sont les paramètres de la loi GG utilisée pour synthétiser les échantillons de  $X(t)$ , nous dirons alors que le processus discret en entrée de la TPOD est un processus AR(1)-GG de paramètres  $\alpha$ ,  $\beta$  et  $\mu$ . Les résultats expérimentaux sont effectués avec  $\mu = 0.5, 0.75, 0.9, 0.95$ . Les spectres des processus correspondant à ces valeurs sont donnés à la figure 2.2. La table 2.1 donne les valeurs  $\gamma(a(\kappa))$  pour les 4 séquences tests. Les résultats expérimentaux obtenus avec le processus AR(1)-GG sont présentés dans la table 3.2.

**Tableau 3.2** — Taux d'acceptation par le test KS de la distribution normale  $\mathcal{N}(0, 1)$  pour les coefficients de la TPOD obtenus aux niveaux de décompositions  $j = 3, 6$  et pour différents chemins. La TPOD est réalisée avec les filtres de Daubechies d'ordres 1 et 7. Le processus d'entrée est AR(1)-GG de paramètres :  $\alpha$  donné par l'Eq. (3.13) et  $\beta = 1.5, 1, 0.5$  et  $\mu = 0.5, 0.75, 0.9$ .

Processus GG coloré, avec $\beta = 1.5$									
Chemin	Filtrage AR(1) : $\mu = 0.5$			Filtrage AR(1) : $\mu = 0.75$			Filtrage AR(1) : $\mu = 0.9$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\kappa_0}$	0%	95%	98%	0%	42%	99%	0%	0%	19%
$\mathcal{P}_{\kappa_1}$	0%	91%	98%	0%	52%	96%	0%	0%	94%
$\mathcal{P}_{\kappa_2}$	0%	95%	88%	0%	37%	86%	0%	0%	91%
$\mathcal{P}_{\kappa_3}$	0%	0%	86%	0%	14%	65%	0%	0%	53%

Processus GG coloré, avec $\beta = 1$									
Chemin	Filtrage AR(1) : $\mu = 0.5$			Filtrage AR(1) : $\mu = 0.75$			Filtrage AR(1) : $\mu = 0.9$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\kappa_0}$	0%	84%	94%	0%	31%	96%	0%	0%	21%
$\mathcal{P}_{\kappa_1}$	0%	94%	96%	0%	67%	93%	0%	0%	92%
$\mathcal{P}_{\kappa_2}$	0%	95%	82%	0%	56%	78%	0%	0%	89%
$\mathcal{P}_{\kappa_3}$	0%	0%	71%	0%	3%	50%	0%	0%	41%

Processus GG coloré, avec $\beta = 0.5$									
Chemin	Filtrage AR(1) : $\mu = 0.5$			Filtrage AR(1) : $\mu = 0.75$			Filtrage AR(1) : $\mu = 0.9$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\kappa_0}$	0%	5%	67%	0%	0%	75%	0%	0%	3%
$\mathcal{P}_{\kappa_1}$	0%	66%	70%	0%	93%	93%	0%	1%	89%
$\mathcal{P}_{\kappa_2}$	0%	10%	4%	0%	71%	46%	0%	0%	29%
$\mathcal{P}_{\kappa_3}$	0%	0%	2%	0%	0%	7%	0%	44%	8%



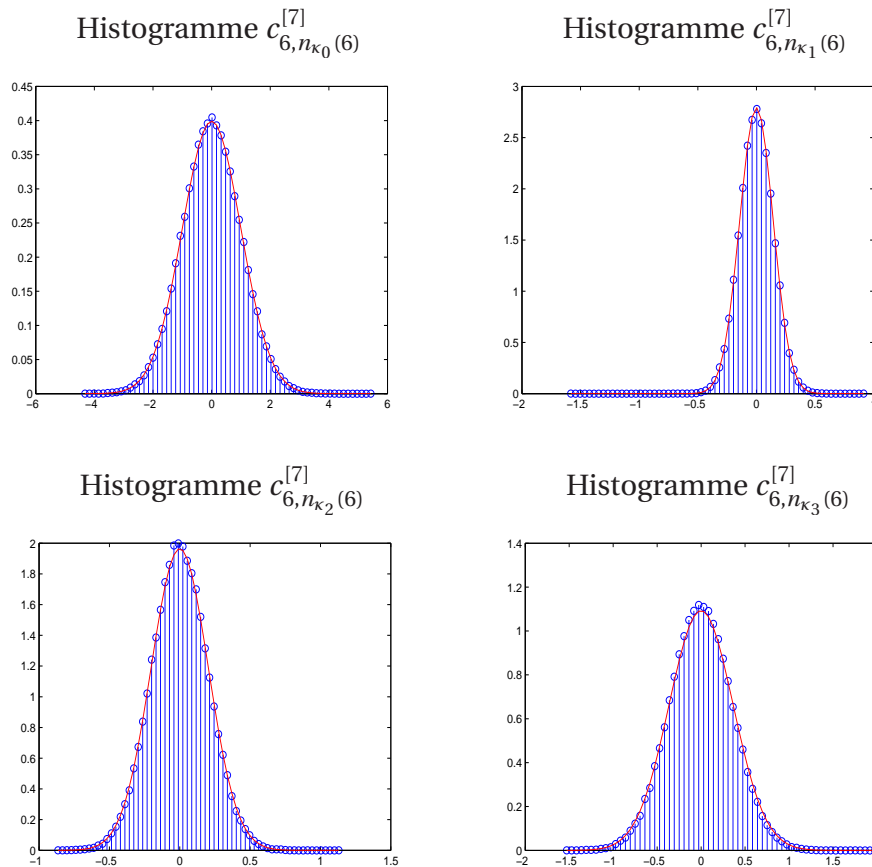
En augmentant le niveau de décomposition de  $j = 3$  à  $j = 6$  lorsque l'ordre des filtres est fixé à  $r = 1$ , le taux d'acceptation par le test KS de la distribution Gaussienne  $\mathcal{N}(0, 1)$  pour la séquence  $(c_{j, n_\kappa(j)}^{[r]}[k] / \sqrt{\gamma(a(\kappa))})_k$  augmente pour la plupart des chemins considérés. D'autre part, lorsque le niveau de décomposition est fixé à  $j = 6$ , il suffit d'augmenter l'ordre  $r$  des filtres de Daubechies pour augmenter le taux d'acceptation par le test KS de la distribution Gaussienne. Si le processus coloré AR(1)-GG  $X(t)$  est de sorte que  $0.5 < \beta \leq 2$  et  $0 < \mu < 0.9$ , la normalité est assez raisonnablement atteinte lorsque le niveau de décomposition est  $j = 6$  et l'ordre des filtres de Daubechies est  $r = 7$ . Les résultats les moins satisfaisants sont obtenus pour des grandes valeurs de  $\mu$  ou pour des petites valeurs de  $\beta$ . Mais il faut noter que lorsque  $\mu$  est grand, le spectre du processus AR(1)-GG tend à devenir singulier autour de la fréquence nulle (voir figure 2.2 pour  $\mu = 0.95$ ) ; d'autre part, lorsque  $\beta$  devient petit, les distributions GGs deviennent singulières en ce sens où leurs formes sont très pointues à l'origine et s'annulent très rapidement (voir figure 3.1 pour  $\beta = 0.5$ ). Cependant, même pour les grandes valeurs de  $\mu$  (lorsque  $0.9 \leq \mu < 1$ ) ou pour les petites valeurs de  $\beta$  (lorsque  $0 < \beta \leq 0.5$ ), le fait d'augmenter le niveau de décomposition et l'ordre des filtres de Daubechies conduit à de meilleurs résultats comme on peut le voir dans la table 3.3.

**Tableau 3.3** — Taux d'acceptation par le test KS de la distribution normale  $\mathcal{N}(0, 1)$  pour les coefficients de la TPOD obtenus aux niveaux de décompositions  $j = 3, 6, 7$  et pour différents chemins. La TPOD est réalisée avec les filtres de Daubechies d'ordres 1, 7 et 20. Le processus d'entrée est AR(1)-GG de paramètres :  $\alpha$  donné par l'Eq. (3.13),  $\beta = 0.5$  et  $\mu = 0.95$ .

Processus GG coloré, avec $\beta = 0.5, \mu = 0.95$					
Chemin	$j = 3$	$j = 6$	$j = 6$	$j = 6$	$j = 7$
	$r = 1$	$r = 1$	$r = 7$	$r = 20$	$r = 20$
$\mathcal{P}_{\kappa_0}$	0%	0%	0%	0%	45%
$\mathcal{P}_{\kappa_1}$	0%	0%	78%	86%	91%
$\mathcal{P}_{\kappa_2}$	0%	0%	2%	30%	79%
$\mathcal{P}_{\kappa_3}$	0%	0%	0%	4%	35%

À partir des tables 3.1, 3.2 et 3.3, on peut donc conclure qu'on observe un taux significatif d'acceptation pour les échantillons  $(c_{j, n_\kappa(j)}^{[r]} / \sqrt{\gamma(a(\kappa))})_k$  de la normalité  $\mathcal{N}(0, 1)$  en augmentant d'abord le niveau de décomposition de lorsque  $r$  est fixé. En outre, on observe une augmentation du taux d'acceptation de la normalité lorsque, pour un niveau de décomposition fixé, on augmente l'ordre des filtres de la TPOD. Ce qui confirme les résultats énoncés par le Théorème 4. A titre illustratif, la figure 3.2 montre les histogrammes des coefficients de la TPOD obtenus à  $j = 6$ , et en utilisant les filtres de Daubechies d'ordre 7. La décomposition concerne les échantillons du processus AR(1)-GG de paramètre  $\alpha$ , donné par l'Eq. (3.13),  $\beta = 1.5$  et  $\mu = 0.75$ . Ces histogrammes sont comparés avec la distribution gaussienne limite.

Comme mentionné un peu plus haut (remarque 2), l'ordre des filtres semble accélérer la convergence pour le chemin des approximations  $\mathcal{P}_{\kappa_0}$ . En effet, pour ce chemin, les filtres d'ordre élevé donnent un plus fort taux d'acceptation de la distribution  $\mathcal{N}(0, \gamma(0))$  que les filtres d'ordre plus faibles. Ce qui peut être observé en comparant, au même niveau de décomposition  $j = 6$ , les taux obtenus pour  $r = 1$  et ceux obtenus pour  $r = 7$  dans les tables 3.1, 3.2 et 3.3 pour  $\mathcal{P}_{\kappa_0}$ .



**Figure 3.2** — Histogrammes des coefficients de la TPOD, au niveau de décomposition 6, et en utilisant les filtres de Daubechies d'ordre 7. La décomposition concerne les échantillons d'un processus AR(1)-GG avec  $\alpha$  donné par l'Eq. (3.13),  $\beta = 1.5$  et  $\mu = 0.75$ . La distribution limite  $\mathcal{N}(0, \gamma(a(\kappa)))$  où  $a(\kappa_0) = 0$  et  $a(\kappa_q) = \pi/2^{q-1}$  pour  $q = 1, 2$  et 3 est représentée en ligne (rouge) continue.

## Conclusion

Dans cette partie, nous nous sommes intéressés aux propriétés statistiques des coefficients intra-bandes de la  $M$ -TPOD d'un processus stationnaire à bande limitée. Nous avons d'abord mis en évidence le fait que la distribution de ces coefficients n'est pas triviale lorsqu'on se donne des filtres quelconques de la  $M$ -TPOD. Ensuite, nous avons montré que pour certaines familles particulières de filtres, on pouvait obtenir le comportement asymptotique des distributions des coefficients. Ces familles particulières de filtres sont celles dont les transformées de Fourier convergent presque partout vers les transformées de Fourier des filtres de Shannon.

Les filtres de Daubechies et ceux de Battle-Lemarié font partie de ces familles de filtres lorsque  $M = 2$ . Plus généralement, et puisque la fonction d'échelle de Daubechies (Battle-Lemarié, respectivement) converge vers la fonction d'échelle de Shannon lorsque le nombre de moments nuls de la fonction d'ondelettes de Daubechies (resp. ordre de la spline de la fonction d'échelle de Battle-Lemarié) augmente [Shen and Strang, 1996] (resp. [Aldroubi, Unser, and Eden, 1992]), on peut espérer que les filtres de la  $M$ -TPOD de Daubechies et de Battle-Lemarié convergent, pour tout  $M \geq 2$ , vers ceux de la  $M$ -TPOD de Shannon. Ce qui cependant n'a pas encore été prouvé formellement, du moins à notre connaissance. Si tel était

le cas, la tendance à la normalité serait accélérée puisque l'Eq. (3.9) suggère qu'augmenter  $M$  est de nature à accroître la vitesse de convergence des cumulants vers 0. Ainsi donc, la synthèse de filtres de  $M$ -TPOD convergeant vers ceux de Shannon est une perspective d'intérêt à la fois théorique et pratique.

La transformée de Karhunen-Loève est souvent qualifiée de discrétisation optimale en ce sens qu'elle permet de transformer certains processus stationnaires au sens large en séquences de données décorréliées. La  $M$ -TPOD peut donc être qualifiée de discrétisation quasi-optimale puisqu'elle permet de transformer des processus stationnaires au sens large en séquences de données quasi-décorréliées (au vu de la qualité de la décorrélation obtenue à la Section 2.3).

D'autre part, les résultats présentés dans la Section 3.2 vont au-delà de la simple décorrélation car ils précisent la nature des distributions asymptotiques de la  $M$ -TPOD de processus stationnaires à bandes limitées. La  $M$ -TPOD permet donc de représenter des processus Hilbertiens stationnaires par des séquences de processus discrets très sensiblement blancs et Gaussiens. Compte tenu de l'importance des processus Gaussiens en statistique et en traitement de signal, ce dernier résultat fait de la  $M$ -TPOD, un outil puissant d'analyse dont les champs d'application ne peuvent que s'élargir davantage.

## **Deuxième partie**

### **Estimation non-paramétrique dans les bases d'ondelettes**



---

# Introduction

Cette partie traite de l'estimation *non-paramétrique* de signaux observés en présence de perturbation (bruit). Le terme “non-paramétrique” désigne les méthodes élaborées sans introduire de connaissances *a priori* sur la distribution du signal étudié. Cette approche non-paramétrique est motivée par le fait que dans la plupart des situations rencontrées en pratique, les distributions des signaux constituant l'observation ne sont pas connues et sont souvent difficiles à estimer avec précision. Notons que dans le cas où l'on connaît les distributions des signaux constituant l'observation, les stratégies *paramétriques* de type *Maximum a Posteriori* sont optimales et donnent de très bonnes performances en pratique.

Lorsqu'on ne dispose pas d'information sur la distribution du signal que l'on cherche à estimer, il est tout de même possible d'estimer ce signal de manière satisfaisante si l'on dispose d'une base qui permet de le distinguer (discriminer) du bruit. L'idée est la suivante. Considérons un bruit additif blanc Gaussien. On sait que la projection d'un tel type de bruit sur une base orthonormale ne change pas la nature de ce bruit. Soit alors une observation constituée d'un signal perturbé par du bruit additif blanc et Gaussien. Si nous disposons d'une base orthonormale capable de “concentrer” l'énergie de ce signal sur un petit nombre de coefficients de grandes amplitudes, alors la projection de l'observation sur une telle base sera assez sensiblement composée d'une petite classe de coefficients relativement grands (donc contenant principalement l'information sur le signal) et une grande classe de petits coefficients dont la nature est presque essentiellement de bruit. L'estimation consisterait alors à récupérer et éventuellement à traiter les grands coefficients, après avoir forcé à zéro, les petits coefficients. Une telle estimation est donc réalisable avec une fonction dite de *seuillage* car elle va dépendre d'un seuil en dessous duquel, tout coefficient sera forcé à zéro.

On dit d'une base capable de concentrer une grande partie de l'énergie d'un signal sur un petit nombre de coefficients de grandes amplitudes qu'elle effectue une *représentation parcimonieuse* (sparse) du signal [Donoho and Johnstone, 1994; Johnstone, 1999]. Cette définition est cependant très peu précise car les termes “petits” et “grands” ne sont pas clairement définis. Nous parlerons donc plutôt de *représentation parcimonieuse au sens large*. Nous dirons alors d'une base qu'elle réalise une *représentation parcimonieuse au sens strict* pour un signal donné si l'énergie de ce signal est “presque entièrement” contenue dans un petit nombre de coefficients de grandes amplitudes. Dans ce dernier cas de parcimonie, presque tous coefficients qualifiés de “petits” sont en fait quasiment nuls ou de très petites amplitudes, et ne contiennent donc pas d'*information significative* sur le signal.

Dans le cas d'une représentation parcimonieuse au sens large, mais non strictement parcimonieuse, il peut être souvent utile de traiter également les petits coefficients, comme nous le

verrons par la suite. Dans ce cas, nous considérerons, non pas une fonction de seuillage, mais une fonction dite d'*atténuation* en ce sens qu'elle ne force pas systématiquement à zéro les petits coefficients, mais leur impose une plus grande atténuation que les grands coefficients. Et puisque la fonction de seuillage peut être vue comme étant une fonction d'atténuation dégénérée (qui opère une atténuation brutale par mise à zéro des petits coefficients et une autre éventuelle atténuation des grands coefficients), nous garderons la terminologie "fonction d'atténuation" pour l'estimation des coefficients *via* la représentation parcimonieuse. Si la fonction d'atténuation inclue un seuillage, nous parlerons alors de "fonction de seuillage".

Pour réaliser l'estimation non-paramétrique décrite un peu plus haut, il faut donc choisir une représentation parcimonieuse et une fonction d'atténuation appropriée. En ce qui concerne le choix de la représentation, on sait [Donoho and Johnstone, 1994] que les bases d'ondelettes satisfont au critère de parcimonie de la représentation pour les signaux réguliers ou réguliers par morceaux. De plus, il a été montré dans [Donoho and Johnstone, 1994] que de simples fonctions de seuillage comme le seuillage *dur* (forcer à zéro ou conserver) et le seuillage *doux* (forcer à zéro ou atténuer d'une valeur égale au seuil) permettent d'obtenir de résultats très satisfaisants. En effet, les performances obtenues pour l'estimation non-paramétrique avec ces fonctions sont comparables aux performances d'un oracle, lorsque ces fonctions de seuillage sont calibrées avec des seuils convenablement choisis [Donoho and Johnstone, 1994]. Un oracle est une procédure élaborée sur la base d'une stratégie d'estimation et dont le but est de recueillir le signal. Ses particularités (comparativement à un estimateur) sont les suivantes :

- il opère sur le signal qu'on cherche à estimer, tandis qu'un estimateur opère sur l'observation ;
- il est optimal au sens de l'erreur quadratique moyenne pour la stratégie d'estimation choisie.

Cette partie traite à la fois du choix des seuils et des fonctions d'atténuation pour l'amélioration de l'estimation non-paramétrique de signaux et d'images. Le modèle de perturbation considéré est le bruit additif blanc et Gaussien. Ce modèle est justifié par le fait que pour une grande classe de processus aléatoires (processus du second ordre, strictement stationnaire, à bande limitée), on peut se ramener assez sensiblement à ce modèle de bruit par utilisation d'une base d'ondelettes constituée de sous-bandes dont le niveau de décomposition et l'ordre des filtres d'analyse sont suffisamment élevés (cf. première partie).

Le Chapitre 4 fait des rappels sur l'estimation non-paramétrique au sens de [Donoho and Johnstone, 1994]. Ensuite, dans le Chapitre 5, nous considérons le problème de la discrimination entre les coefficients du signal et ceux du bruit, à partir d'une fonction de seuillage classique, la fonction de seuillage doux. Nous traitons le choix du seuil pour la séparation entre les coefficients qui seront attribués au bruit (et seront donc forcés à zéro), et ceux qui seront attribués au signal. Ce problème peut être vu comme un problème de détection de signaux dans un environnement bruité. Pour résoudre ce problème, nous adaptons les paramètres d'une méthode de détection non-paramétrique [Pastor, Gay, and Gronenboom, 2002]. Nous montrons ensuite, par des tests basés sur la fonction de seuillage doux (connue pour ses propriétés de "lissage" et d'"adaptation", [Donoho, 1995]), que les seuils obtenus permettent d'améliorer sensiblement les résultats de l'estimation non-paramétrique. Cette analyse révèle le fait que l'estimation par seuillage est plus adaptée aux signaux admettant une représentation parcimonieuse au sens strict.

Enfin, et compte tenu du fait que le modèle de parcimonie au sens strict n'est pas souvent

vérifié en pratique, nous introduisons, dans le Chapitre 6, une nouvelle famille de fonctions qui opèrent une atténuation sigmoïdale sur les coefficients d'ondelettes. Cette classe de fonction évite la mise à zéro systématique des petits coefficients. Les éléments de cette classe sont appelés les fonctions SSBS (pour Smooth Sigmoid Based Shrinkage). Une fonction SSBS est une fonction très régulière, ce qui possède l'avantage d'éviter les artefacts de reconstruction. Elle dépend de paramètres qui la rendent très flexible de sorte que l'on peut contrôler le degré d'atténuation imposé aux coefficients. Le choix et l'optimisation des paramètres des SSBS sont également discutés pour cette nouvelle classe de fonction.





## Estimation non-paramétrique - Rappels et état de l'art

Soit  $\mathbf{y} = \{y_i\}_{1 \leq i \leq N}$  une séquence de données (observation) :  $y_i = f(t_i) + e_i$ ,  $i = 1, 2, \dots, N$ , où  $f$  est une fonction inconnue, les variables aléatoires  $\{e_i\}_{1 \leq i \leq N}$  sont indépendantes et identiquement distribuées (iid). On suppose que chaque  $e_i$  est de distribution Gaussienne, de moyenne nulle et de variance  $\sigma^2$  :  $e_i \sim \mathcal{N}(0, \sigma^2)$  pour tout  $i = 1, 2, \dots, N$ .

Le problème traité dans ce chapitre concerne l'estimation non-paramétrique du signal  $\{f(t_i)\}_{1 \leq i \leq N}$  selon l'approche développée dans [Donoho and Johnstone, 1994]. Pour estimer  $f(\cdot)$ , on projette d'abord  $\mathbf{y}$  sur une base orthonormale représentée par une matrice unitaire  $\mathcal{W}$ . Le résultat de cette projection est une séquence de coefficients

$$c_i = d_i + \epsilon_i, \quad i = 1, 2, \dots, N, \quad (4.1)$$

avec  $\mathbf{c} = \{c_i\}_{1 \leq i \leq N} = \mathcal{W} \mathbf{y}$ ,  $\mathbf{d} = \{d_i\}_{1 \leq i \leq N} = \mathcal{W} \mathbf{f}$ ,  $\mathbf{f} = \{f(t_i)\}_{1 \leq i \leq N}$  et  $\boldsymbol{\epsilon} = \{\epsilon_i\}_{1 \leq i \leq N} = \mathcal{W} \mathbf{e}$ ,  $\mathbf{e} = \{e_i\}_{1 \leq i \leq N}$ . Les variables aléatoires  $\{\epsilon_i\}_{1 \leq i \leq N}$  sont iid et  $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ .

Si l'on note  $\hat{\mathbf{d}} = \{\delta(c_i)\}_{1 \leq i \leq N}$  le résultat du traitement effectué sur les coefficients  $\{c_i\}_{1 \leq i \leq N}$  par une fonction  $\delta(\cdot)$ , l'estimée de  $\mathbf{f}$  est alors  $\hat{\mathbf{f}} = \mathcal{W}^\top \hat{\mathbf{d}}$ , où  $\mathcal{W}^\top$  est la transposée et donc l'inverse de  $\mathcal{W}$  (supposée être une matrice unitaire).

Soit  $\hat{\mathbf{f}}$  une estimée de  $\mathbf{f}$  calculée en appliquant le traitement  $\delta$ . Le risque ou fonction de coût utilisé pour mesurer la qualité de l'estimation est l'Erreur Quadratique Moyenne (EQM). Puisque la transformée  $\mathcal{W}$  est orthonormale, ce coût est

$$r(\mathbf{d}, \hat{\mathbf{d}}) = \frac{1}{N} \mathbb{E} \|\mathbf{d} - \hat{\mathbf{d}}\|^2 = \frac{1}{N} \sum_{i=1}^N \mathbb{E} (d_i - \delta(c_i))^2.$$

Nous considérons, à titre de comparaison, le coût suivant associé à un oracle et appelé *risque d'oracle*.

$$r_0(\mathbf{d}) = \frac{1}{N} \sum_{i=1}^N \min(d_i^2, \sigma^2). \quad (4.2)$$

Les notions d'oracles et de risque d'oracle sont détaillées dans [Donoho and Johnstone, 1994].

On suppose que la transformée  $\mathcal{W}$  représente le signal  $f$  de manière parcimonieuse. Ainsi, seul un petit nombre de coefficients  $d_i$ ,  $i = 1, 2, \dots, N$  admettent de grandes amplitudes, et en ce sens, caractérisent le signal. Il est alors relativement plus simple d'estimer  $\mathbf{d}$  que d'estimer  $\mathbf{f}$ . La section suivante donne les procédures couramment utilisées pour cette estimation lorsque les représentations parcimonieuses concernées sont les transformées par ondelettes (TOD, M-TPOD).

## 4.1 Parcimonie de la représentation par ondelettes

Les transformées par ondelettes sont parcimonieuses pour les signaux réguliers par morceaux [Donoho and Johnstone, 1994], [Johnstone, 1999], comme on peut le voir sur les figures 4.1 et 4.2. Ces figures représentent des signaux mono et bi-dimensionnels (images), ainsi que leurs TODs basées une ondelette “Symlet” d’ordre 8 (sym8) ; la décomposition concerne les niveaux de résolution  $j = 1, 2, \dots, J$ , avec  $J = 6, 10$  pour les signaux et  $J = 4$  pour les images. Notons qu’on a sensiblement le même type de parcimonie pour une TOD que pour une TPOD (et pour une  $M$ -TPOD par extension). En effet, les coefficients d’approximation, qui constituent la majeure partie des coefficients de grandes amplitudes, sont communs à la TOD et à la TPOD. D’autre part, la TPOD s’obtient à partir de la TOD en décomposant récursivement les coefficients de détails, ce qui ne change pas considérablement la nature parcimonieuse des résultats obtenus.

Par la suite, nous supposons que la transformée représentée par la matrice  $\mathcal{W}$  est une transformée en ondelettes. La parcimonie induite par les transformées en ondelettes permet de discriminer les coefficients du signal de ceux du bruit additif blanc Gaussien. En effet, ce type de bruit se distribue sur les petits coefficients et affecte très peu les grands coefficients, comme on peut le voir sur la figure 4.3. Compte tenu de cette nature parcimonieuse, une première étape dans la procédure d’estimation des coefficients du signal consiste en la mise à zéro des petits coefficients puisqu’il est *a priori* difficile de distinguer la contribution du signal dans ces coefficients. Cette étape de mise à zéro des petits coefficients se réalise à l’aide d’une fonction de seuillage.

Diverses fonctions de seuillage ont été proposées dans la littérature. Parmi ces fonctions de seuillages, les plus populaires sont données par la figure 4.4. En plus des fonctions de seuillage dur et doux proposées par [Donoho and Johnstone, 1994], nous avons également représenté sur cette figure, les fonctions dites du Non-Negative Garrote (NNG) de [Breiman, 1995] (proposée pour le seuillage dans le domaine des ondelettes par [Gao, 1998]) et du SCAD (pour Smoothly Clipped Absolute Deviation) introduite dans [Antoniadis and Fan, 2001]. Ces deux dernières fonctions réalisent un certain compromis entre le seuillage dur et le seuillage doux, comme on peut le voir sur la figure 4.4.

Ces fonctions de seuillage classiques dépendent toutes d’un seuil  $\lambda$ , de sorte que tout coefficient  $c_i$ ,  $1 \leq i \leq N$  dont l’amplitude est inférieure à  $\lambda$  est forcé à 0. La section suivante donne des résultats concernant la fonction de seuillage doux.

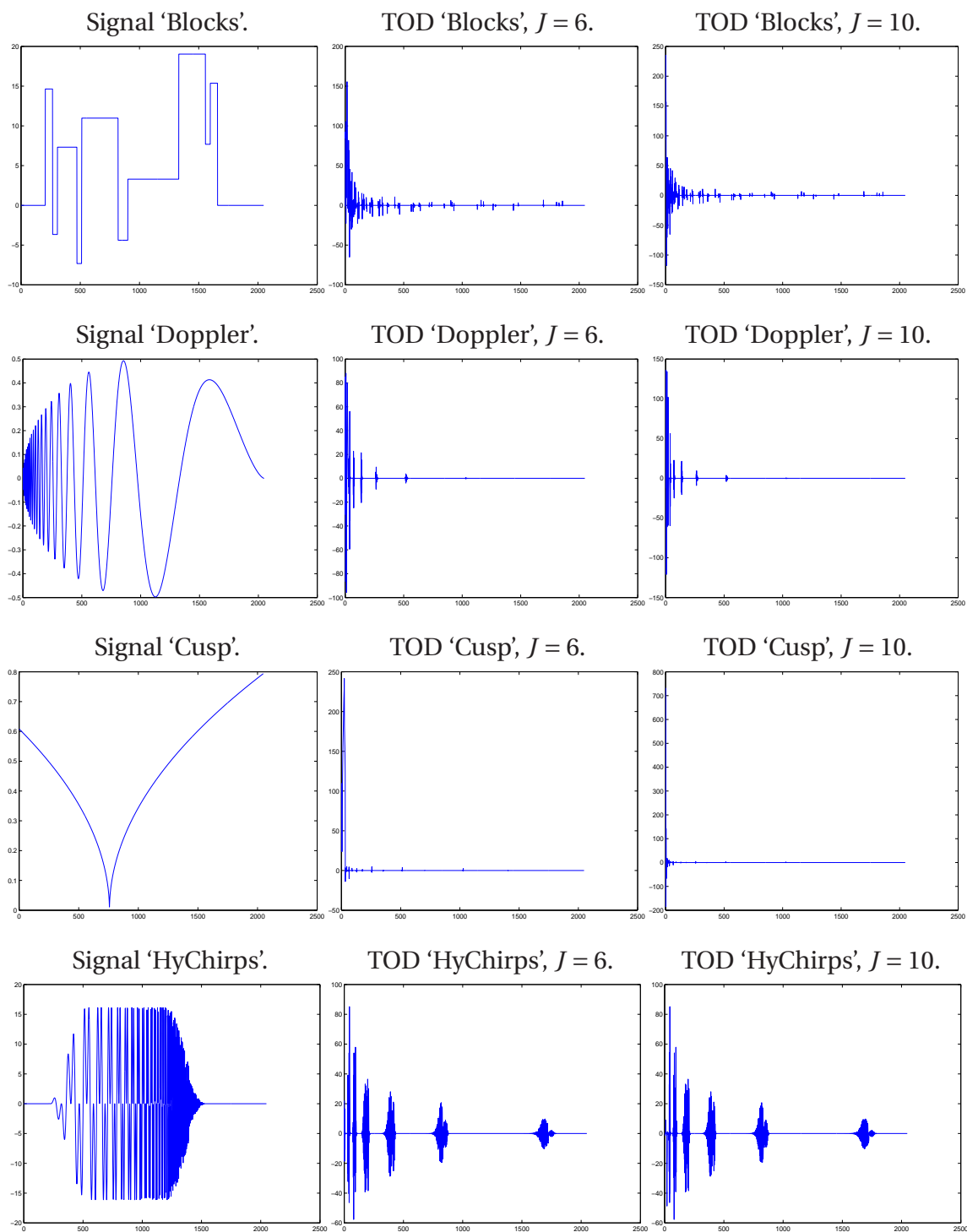
## 4.2 Estimation non-paramétrique par seuillage doux

Dans la suite de ce chapitre, nous considérons la fonction de seuillage doux, qui est très appréciée pour ses propriétés de lissage et d’adaptation [Donoho, 1995]. Son expression est donnée par

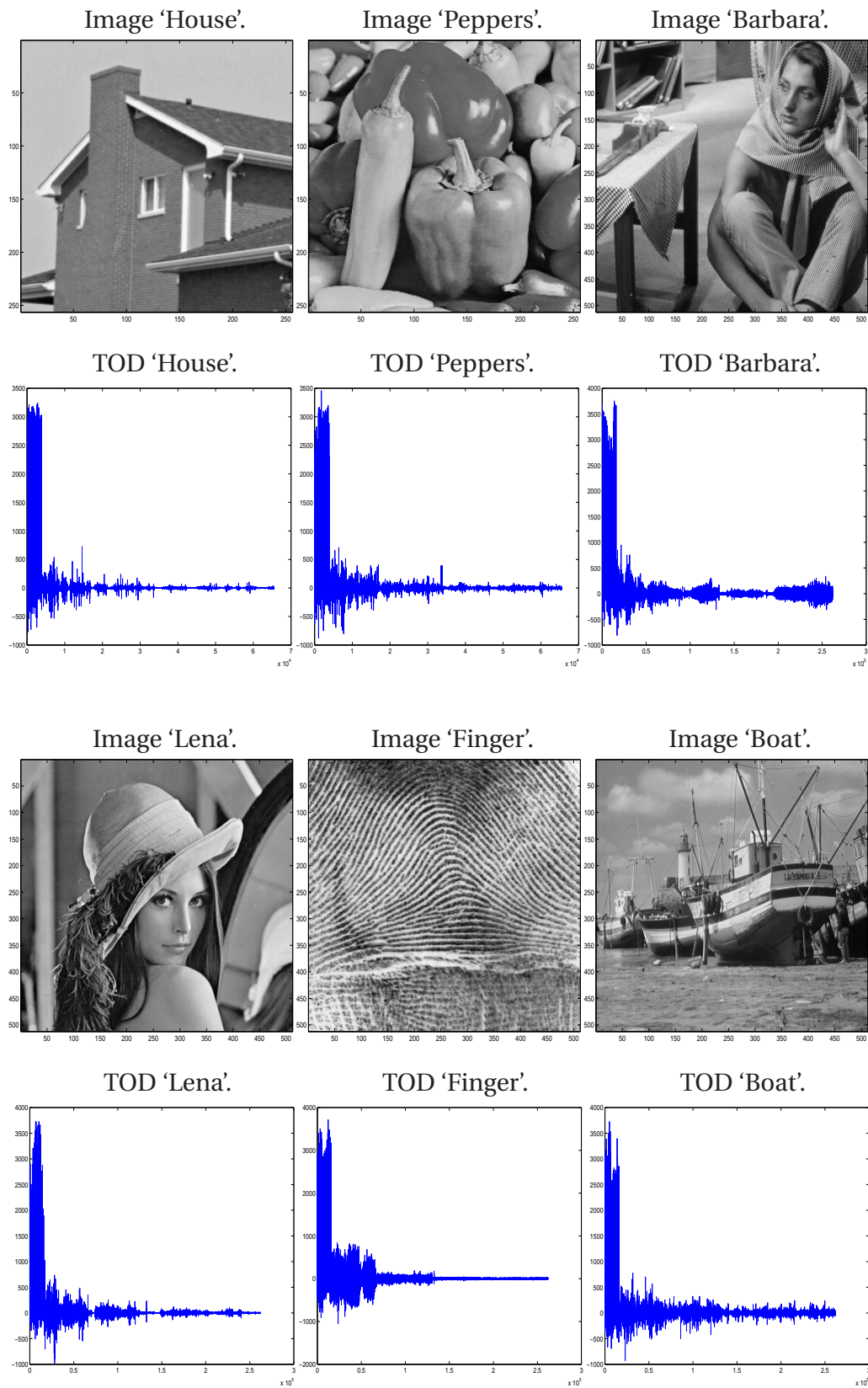
$$\delta_\lambda(x) = \begin{cases} x - \operatorname{sgn}(x)\lambda & \text{si } |x| \geq \lambda, \\ 0 & \text{sinon,} \end{cases} \quad (4.3)$$

où  $\operatorname{sgn}(x) = 1$  (resp.  $-1$ ) si  $x \geq 0$  (resp.  $x < 0$ ).

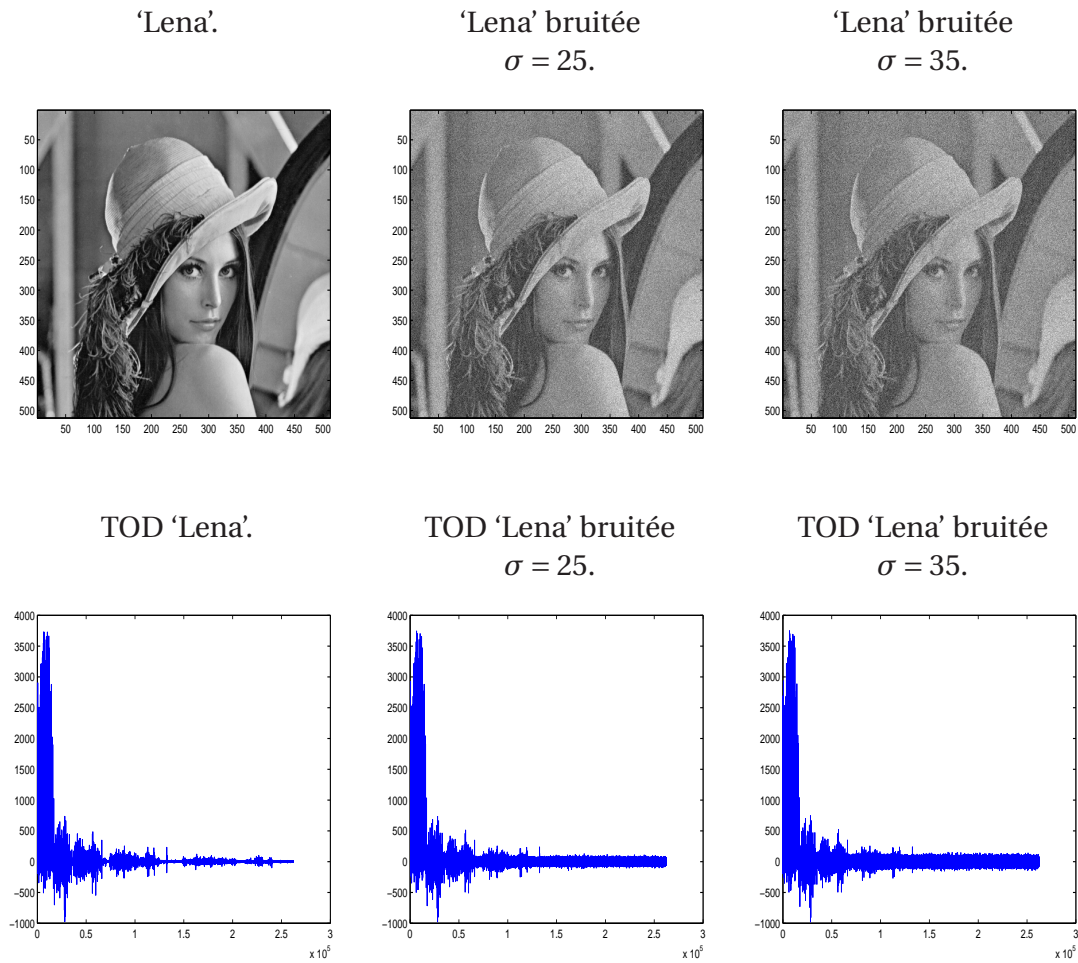
Le problème du choix des seuils (pour la distinction entre la contribution du signal et



**Figure 4.1** — Exemples de signaux tests, avec leurs représentations par TOD. La TOD concerne les niveaux de décomposition  $j = 1, 2, \dots, J$  où  $J = 6, 10$ .



**Figure 4.2** — Quelques images naturelles, ainsi que leurs représentations par TOD avec une ondelette 'sym8' et  $J = 4$ .



**Figure 4.3** — Représentations par TOD avec une ondelette ‘sym8’ de l’image originale et des images bruitées de ‘Lena’, pour  $J = 4$ .

celle du bruit) a été longuement discuté dans [Donoho and Johnstone, 1994; Donoho, 1995; Donoho and Johnstone, 1995]. Rappelons les expressions des seuils universels et minimax, pour lesquels nous disposons d’une majoration du risque lorsque le seuillage doux est ajusté avec ces seuils. Considérons l’Eq. (4.1). Puisque  $\epsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ , on obtient de [Berman, 1992, Eqs. (9.2.1), (9.2.2), Section 9.2, p. 187] (voir aussi [Mallat, 1999, p. 454], [Serfling, 1980, Section 2.4.4, p. 91]), que

$$\lim_{N \rightarrow +\infty} \mathbb{P} \left[ \lambda_u(N) - \frac{\sigma \ln \ln N}{\ln N} \leq \max \{ |\epsilon_i|, 1 \leq i \leq N \} \leq \lambda_u(N) \right] = 1, \quad (4.4)$$

où  $\lambda_u(N) = \sigma \sqrt{2 \ln N}$ . Ainsi, le maximum en amplitude de  $\{\epsilon_i\}_{1 \leq i \leq N}$  a une forte probabilité d’être proche de  $\lambda_u(N)$  lorsque  $N$  est grand. Le seuil  $\lambda_u(N)$  est le *seuil universel*. D’après [Donoho and Johnstone, 1994, Theorem 1], le risque  $r_{\lambda_u(N)}(\mathbf{d}, \hat{\mathbf{d}})$  de l’estimation par seuillage doux de  $\mathbf{d}$  avec le seuil universel  $\lambda_u(N)$  est tel que

$$r_{\lambda_u(N)}(\mathbf{d}, \hat{\mathbf{d}}) \leq (1 + 2 \ln N) (N^{-1} \sigma^2 + r_0(\mathbf{d})). \quad (4.5)$$

Le *seuil minimax*  $\lambda_m(N)$  est défini comme étant la plus grande valeur  $\lambda$  parmi les valeurs

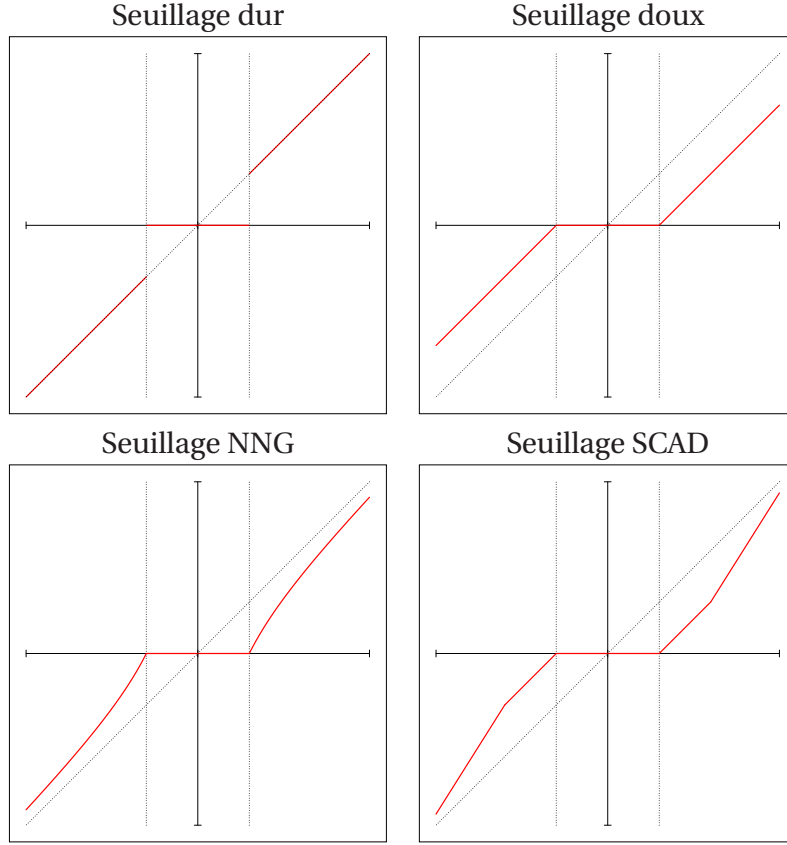


Figure 4.4 — Graphes des fonctions de seuillage dur, doux, NNG et SCAD.

qui atteignent la borne du risque minimax  $\Lambda(N)$  définie par

$$\Lambda(N) = \inf_{\lambda > 0} \sup_{\mu \in \mathbb{R}} \frac{r_{\lambda}(\mu, \hat{\mu})}{N^{-1} + r_0(\mu)}. \quad (4.6)$$

De [Donoho and Johnstone, 1994, Theorem 2], on obtient que le risque  $r_{\lambda_m(N)}(\mathbf{d}, \hat{\mathbf{d}})$  de l'estimation par seuillage doux de  $\mathbf{d}$  avec le seuil minimax  $\lambda_m(N)$  satisfait l'inégalité

$$r_{\lambda_m(N)}(\mathbf{d}, \hat{\mathbf{d}}) \leq \Lambda(N) (N^{-1} \sigma^2 + r_0(\mathbf{d})), \quad (4.7)$$

avec  $\Lambda(N) \leq 1 + 2 \ln N$  et  $\Lambda(N) \stackrel{N \rightarrow \infty}{\sim} 2 \ln N$ .

**Remarque 3** D'après les inégalités données par les Eqs. (4.5) et (4.7), la borne supérieure du risque  $r_{\lambda}(\mathbf{d}, \hat{\mathbf{d}})$  de l'estimation par seuillage doux est de l'ordre de  $2r_0(\mathbf{d}) \ln N$  lorsque  $N$  tend vers l'infini,  $\lambda$  étant l'un quelconque des seuils universel ou minimax. La quantité  $2r_0(\mathbf{d}) \ln N$  est l'ordre minimal que l'on puisse espérer pour la borne supérieure du risque sur l'ensemble des estimateurs diagonaux, dont le seuillage doux en est un élément [Donoho and Johnstone, 1994, Theorem 3]. Ce qui justifie de l'intérêt des seuils universel et minimax en estimation par seuillage doux. Comme nous le verrons dans le chapitre suivant, il est possible, en réduisant la classe d'appartenance des signaux considérés, de déduire des seuils qui admettent une borne supérieure plus petite que le facteur  $2r_0(\mathbf{d}) \ln N$  pour le risque.

Le lemme suivant [Atto, Pastor, and Mercier, 2008a] est très utile pour le calcul des bornes supérieures du risque de l'estimation par seuillage doux.

**Lemme 1** *Etant donné le modèle décrit par l'Eq. (4.1), considérons l'estimation de  $\mathbf{d}$  par seuillage doux avec un seuil (valeur réelle positive)  $\lambda$ .*

*Le risque  $r_\lambda$  de cette estimation est tel que*

$$r_\lambda(\mathbf{d}, \hat{\mathbf{d}}) \leq (1 + \lambda^2/\sigma^2) \times \left( \sigma^2 e^{-\lambda^2/2\sigma^2} + r_0(\mathbf{d}) \right). \quad (4.8)$$

**Preuve :** Voir [Atto, Pastor, and Mercier, 2008a]. ■





## Les seuils de détection pour l'estimation non-paramétrique

Dans ce chapitre, nous présentons les seuils de détection pour l'estimation non-paramétrique. Les seuils de détection ont été définis dans un contexte général de détection non-paramétrique dans [Pastor, Gay, and Gronenboom, 2002], et ont été adaptés à l'estimation non-paramétrique par seuillage dans le domaine des ondelettes dans [Atto, Pastor, and Mercier, 2008a]. Nous considérons l'estimation non-paramétrique par la fonction de seuillage doux pour les raisons évoquées à la Section 4.2. Les performances réalisées avec les seuils de détection seront comparées avec celles obtenues par les seuils universels et minimax. La comparaison est effectuée tant en terme de borne sur le risque qu'en terme de qualité visuelle des signaux et images débruités.

Nous montrerons en particulier que la borne supérieure du risque de l'estimation par seuillage doux avec les seuils de détection se comporte comme  $r_0(\mathbf{d}) \ln N$ , ou encore  $r_0(\mathbf{d}) \ln N/2$ , lorsque  $N$  est grand, pour certaines classes de signaux. Les classes concernées englobent de nombreux signaux rencontrés en traitement de signal et d'images.

### 5.1 Les seuils de détection

Considérons le problème de l'estimation de  $\mathbf{d}$  à partir de l'observation des coefficients  $c_i$ ,  $i = 1, 2, \dots, N$  issus de la représentation parcimonieuse  $\mathcal{W}$  (modèle donné par l'Eq. (4.1)).

Le rôle primordial des seuils dans les fonctions de seuillage est de décider des coefficients qui seront forcés à zéros car ils ne contiennent pas - ou contiennent trop peu - d'information sur le signal. Posons  $\mathcal{T}_\lambda(\cdot)$  le *test à seuil* de hauteur  $\lambda$  (valeur réelle positive) défini pour tout  $x$  réel par

$$\mathcal{T}_\lambda(x) = \begin{cases} 1 & \text{si } |x| \geq \lambda, \\ 0 & \text{sinon.} \end{cases} \quad (5.1)$$

Alors pour tout coefficient  $c_i$ ,  $i = 1, 2, \dots, N$ , et lorsque la fonction de seuillage utilisée est celle du seuillage doux  $\delta_\lambda$ , on a  $\delta_\lambda(c_i) = \mathcal{T}_\lambda(c_i)(c_i - \text{sgn}(c_i)\lambda)$ .

Soit  $V(\rho, p)$  la fonction définie pour tout réel non-négatif  $\rho$  et tout  $0 \leq p \leq 1$  par

$$V(\rho, p) = p [F(\rho + \xi(\rho, p)) - F(\rho - \xi(\rho, p))] + 2(1 - p) (1 - F(\xi(\rho, p))), \quad (5.2)$$

où  $F$  est la fonction de répartition de la distribution normale  $\mathcal{N}(0, 1)$  et

$$\xi(\rho, p) = \frac{\rho}{2} + \frac{1}{\rho} \left[ \ln \frac{1-p}{p} + \ln \left( 1 + \sqrt{1 - \frac{p^2}{(1-p)^2} e^{-\rho^2}} \right) \right]. \quad (5.3)$$

**Proposition 2** *Considérons le test binaire à deux hypothèses*

$$\begin{cases} \mathcal{H}_0 : Z \sim \mathcal{N}(0, \sigma^2) \\ \mathcal{H}_1 : Z = S + X, S \neq 0 \text{ (p-s)}, |S| \geq a \geq 0 \text{ (p-s)}, X \sim \mathcal{N}(0, \sigma^2), \end{cases} \quad (5.4)$$

où  $U, S, X$  désignent des variables aléatoires réelles,  $S$  et  $X$  sont indépendantes ; (p-s) signifie que la propriété est vraie “presque sûrement”.

Soit  $\mathcal{L}$  le test de Bayes à minimum de probabilité d'erreur pour la décision selon les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$  donnée par le système (5.4).

Si la probabilité a priori d'occurrence de l'hypothèse  $\mathcal{H}_1$  est inférieure ou égale à une certaine valeur  $p^*$ , avec  $p^* \leq 1/2$ , alors  $V(a/\sigma, p^*)$  est une borne supérieure de la probabilité d'erreur du test de Bayes  $\mathcal{L}$  et du test à seuil  $\mathcal{T}_{\sigma\xi(a/\sigma, p^*)}$  de hauteur  $\sigma\xi(a/\sigma, p^*)$ . La borne  $V(a/\sigma, p^*)$  est atteinte par les deux tests  $\mathcal{L}$  et  $\mathcal{T}_{\sigma\xi(a/\sigma, p^*)}$  lorsque  $|S| = a$  (p-s), avec  $P[S = a] = P[S = -a] = 1/2$  et la probabilité d'occurrence de l'hypothèse  $\mathcal{H}_1$  est  $p^*$ .

**Preuve :** Voir [Atto, Pastor, and Mercier, 2008a]. ■

La Proposition 2 est valide si les variables aléatoires sont remplacées par des vecteurs aléatoires de dimension  $n$  quelconque et les valeurs absolues remplacées par la norme Euclidienne sur  $\mathbb{R}^n$ . Le résultat ainsi obtenu est une simple extension de [Pastor, Gay, and Gronenboom, 2002, Theorem VII.1].

Si la distribution du signal étudié est connue, ce qui est rarement le cas, le test  $\mathcal{L}$  peut être utilisé pour décider si les données observées contiennent de l'information significative ou non, et cette décision est optimale en ce sens où  $\mathcal{L}$  est le test à minimum de probabilité d'erreur. Dans le cas où l'on ne connaît pas la distribution du signal et l'on ne dispose pas de suffisamment d'information pour estimer cette distribution avec précision, le test  $\mathcal{L}$  ne peut pas être appliqué. Mais la Proposition 2 fournit un test non-paramétrique à seuil  $\mathcal{T}_{\sigma\xi(a/\sigma, p^*)}$ , de seuil  $\lambda_D(a, p^*) = \sigma\xi(a/\sigma, p^*)$  qui garantit la même borne supérieure de la probabilité d'erreur que  $\mathcal{L}$ .

Le seuil  $\lambda_D(a, p^*)$  dépend de deux paramètres qui permettent de décrire le degré de parcimonie du signal étudié [Pastor and Atto, 2008]. En effet, le signal  $S$  est supposé d'amplitude supérieure ou égale à une certaine valeur  $a$ , et  $p^*$  mesure une borne sur la fréquence d'apparition du signal dans l'observation. Notons que  $a$  peut être assimilé à une borne inférieure sur les amplitudes des coefficients significatifs du signal, ce qui correspond à l'amplitude minimale à partir de laquelle on veut décider que la donnée observée contient de l'information significative. Le fait que  $p^* \leq 1/2$  garantit que les coefficients significatifs du signal sont moins souvent

présents qu'absent, ce qui est compatible avec l'hypothèse de parcimonie. Lorsque  $a$  est grand et  $p^*$  très petit (devant  $1/2$ ), on retrouve le modèle de parcimonie au sens strict décrit dans l'introduction de cette seconde partie du document. Le seuil de détection peut donc s'adapter à la parcimonie de la représentation du signal analysé par un choix très judicieux de  $a$  et  $p^*$ , lorsqu'on peut estimer ces valeurs.

Revenons aux coefficients  $\{c_i\}_{1 \leq i \leq N}$  issus de la projection de l'observée  $\mathbf{y}$  via la transformée  $\mathcal{W}$ . Nous considérons que tout coefficient  $c_i$ , pour  $i = 1, 2, \dots, N$ , suit un modèle d'hypothèse binaire dans lequel l'hypothèse nulle est que  $c_i$  contient seulement du bruit, de sorte que  $c_i = \epsilon_i$ , et que l'hypothèse alternative est que  $c_i$  est la somme de signal et du bruit, de sorte que  $c_i = d_i + \epsilon_i$  avec  $d_i \neq 0$ . En pratique, lorsque l'on ne connaît pas  $a$  et  $p^*$  de façon précise et qu'on ne veut pas imposer de contraintes sur le degré de parcimonie de la transformée  $\mathcal{W}$ , on peut considérer les valeurs  $a = \lambda_u(N) = \sigma\sqrt{2\ln N}$  et  $p^* = 1/2$ , hypothèses raisonnables pour couvrir le cas de parcimonie au sens large pour les raisons suivantes :

- (1) On sait que le maximum des coefficients du bruit a une forte probabilité d'être proche de la valeur  $\lambda_u(N)$  lorsque  $N$  est grand. Aussi, toute valeur  $d_i$  telle que  $|d_i| < \lambda_u(N)$  peut se confondre avec une valeur du bruit. On peut donc décider que les coefficients significatifs du signal sont les  $d_i$  pour lesquels  $|d_i| \geq \lambda_u(N)$ .
- (2) D'autre part, puisque la transformée  $\mathcal{W}$  est parcimonieuse, il est légitime de considérer que la proportion des coefficients significatifs, qui joue un rôle similaire à celui de  $p^*$ , est telle que  $p^* \leq 1/2$ . Si on ne dispose pas de précision sur le degré de parcimonie de la représentation, le choix le moins restrictif serait donc de considérer la borne supérieure  $p^* = 1/2$ .

Avec ces hypothèses, et compte tenu de la Proposition 2 et l'Eq. (5.3), l'expression du seuil de détection  $\lambda_D(N) = \lambda_D(\lambda_u(N), 1/2)$  est donnée par

$$\begin{aligned}\lambda_D(N) &= \sigma \xi(\lambda_u(N)/\sigma, 1/2) \\ &= \sigma \sqrt{\ln N/2} + \sigma \ln \left( 1 + \sqrt{1 - 1/N^2} \right) / \sqrt{2 \ln N}.\end{aligned}\tag{5.5}$$

Le test à seuil  $\mathcal{T}_{\lambda_D(N)}$  accepte ou rejette l'hypothèse nulle avec une probabilité d'erreur inférieure ou égale à  $V(\sqrt{2 \ln N})$ , qui est une fonction décroissante de  $N$  (cf. table 5.1).

**Tableau 5.1** — Bornes supérieures  $V(\sqrt{2 \ln N})$  de la probabilité d'erreur du test à seuil  $\mathcal{T}_{\lambda_D(N)}$ .

$N$	2	4	8	16	32
$V(\sqrt{2 \ln N})$	0.3645	0.2743	0.2110	0.1648	0.1302
$N$	64	128	512	1024	2048
$V(\sqrt{2 \ln N})$	0.1036	0.0830	0.0540	0.0437	0.0356
$N$	4096	8192	16384	32768	65536
$V(\sqrt{2 \ln N})$	0.0290	0.0236	0.0193	0.0158	0.0130

La table 5.2 donne les valeurs de  $\lambda_D(N)$ ,  $\lambda_m(N)$ , et  $\lambda_u(N)$  pour certaines valeurs de  $N$ . Cette table montre que pour les petites valeurs de  $N$ , le seuil  $\lambda_D(N)$  est proche du seuil minimax, et pour les grandes valeurs de  $N$  (supérieures ou égales à 2048), le seuil  $\lambda_D(N)$  est de l'ordre de  $\lambda_u(N)/2$ .

**Tableau 5.2** — Seuils de détection, minimax, et universel pour différentes valeurs de  $N$ .

$N$	128	256	512	1024	2048	4096	8192
$\lambda_D(N)$	1.78	1.87	1.96	2.05	2.13	2.21	2.29
$\lambda_m(N)$	1.67	1.86	2.05	2.23	2.40	2.58	2.74
$\lambda_u(N)$	3.12	3.33	3.53	3.72	3.91	4.08	4.25

Les équations (4.5) et (4.7) donnent les bornes des risques de l'estimation par seuillage doux avec les seuils universels et minimax respectivement. La proposition suivante [Atto, Pastor, and Mercier, 2008a] donne le risque de l'estimation par seuillage doux avec le seuil  $\lambda_D(N)$ .

**Proposition 3** *Etant donné le modèle décrit par l'Eq. (4.1), considérons l'estimation de  $\mathbf{d}$  par seuillage doux avec le seuil de détection  $\lambda_D(N)$ . Pour  $N \geq 2$ , le risque  $r_{\lambda_D(N)}$  de cette estimation satisfait à l'inégalité*

$$r_{\lambda_D(N)}(\mathbf{d}, \hat{\mathbf{d}}) \leq (\ln N/2 + \eta(N)) (\sigma^2 \zeta(N) + r_0(\mathbf{d})), \quad (5.6)$$

avec

$$\eta(N) = 1 + \ln \left( 1 + \sqrt{1 - 1/N^2} \right) + \ln^2 \left( 1 + \sqrt{1 - 1/N^2} \right) / 2 \ln N, \quad (5.7)$$

et

$$\zeta(N) = N^{-1/4} \left( 1 + \sqrt{1 - 1/N^2} \right)^{-1/2} \times e^{-\ln^2 \left( 1 + \sqrt{1 - 1/N^2} \right) / 4 \ln N}. \quad (5.8)$$

Les bornes des risques de l'estimation par seuillage doux avec les seuils universels et minimax sont de l'ordre de  $2r_0(\mathbf{d}) \ln N$  lorsque  $N$  tend vers l'infini (voir les Eqs. (4.5) et (4.7)). Par la suite, nous allons montrer que la borne supérieure du risque de l'estimation par seuillage doux avec le seuil de détection  $\lambda_D(N)$  est de l'ordre de 2 à 4 fois plus petite que  $2r_0(\mathbf{d}) \ln N$  lorsque  $N$  tend vers l'infini, et cela lorsque  $\mathbf{d}$  appartient à une certaine classe de signaux. Notons que d'après [Donoho and Johnstone, 1994, Theorem 3], la quantité  $2r_0(\mathbf{d}) \ln N$  est l'ordre minimal que l'on peut espérer pour la borne supérieure du risque par utilisation des estimateurs diagonaux tel que le seuillage doux.

Considérons le sous-ensemble

$$\Theta_N = \{ \mathbf{d} = \{d_i\}_{1 \leq i \leq N} \in \mathbb{R}^N : r_0(\mathbf{d}) \geq \sigma^2 \zeta(N) \}$$

de  $\mathbb{R}^N$ . Les éléments de ce sous-ensemble peuvent être vus comme des séquences de coefficients issus de la projection *via*  $\mathcal{W}$  pour certaines classes de signaux. La classe  $\Theta_N$  sera donc qualifiée de “classe de signaux” dans ce qui suit.

Si  $\mathbf{d}$  appartient à  $\Theta_N$ , la borne supérieure donnée par l'Eq. (5.6) pour le risque de l'estimation par seuillage doux avec le seuil de détection se comporte comme  $r_0(\mathbf{d}) \ln N$  lorsque  $N$  tend vers l'infini. Cela vient du fait que  $\lim_{N \rightarrow \infty} \eta(N) = 1 + \ln 2$  et  $\sigma^2 \zeta(N) + r_0(\mathbf{d}) \leq 2r_0(\mathbf{d})$  lorsque  $\mathbf{d} \in \Theta_N$ .

En outre, pour tout élément  $\mathbf{d}$  de  $\Theta_N$  tel que  $r_0(\mathbf{d}) \gg \sigma^2 \zeta(N)$ , l'ordre de la borne supérieure sur le risque  $r_{\lambda_D(N)}(\mathbf{d}, \hat{\mathbf{d}})$  est dans ce cas  $r_0(\mathbf{d}) \ln N/2$  lorsque  $N$  tend vers l'infini. En

effet, puisque  $\lim_{N \rightarrow \infty} \zeta(N) = 0$ , et si  $r_0(\mathbf{d})$  est très grand en comparaison de  $\sigma^2 \zeta(N)$ , alors  $\sigma^2 \zeta(N) + r_0(\mathbf{d}) \sim r_0(\mathbf{d})$  pour les valeurs suffisamment grandes de  $N$ .

Ainsi, pour l'estimation d'un élément de  $\Theta_N$  par seuillage doux, et lorsque  $N$  est grand, le seuil de détection conduit à une borne supérieure du risque de l'ordre de 2 à 4 fois plus petite que celles obtenues avec les seuils universel et minimax qui sont de l'ordre de  $2r_0(\mathbf{d}) \ln N$  sur  $\Theta_N$ . Le fait d'obtenir une borne supérieure du risque plus petite que la quantité optimale  $2r_0(\mathbf{d}) \ln N$  peut paraître surprenant. Mais il faut noter à ce niveau qu'il n'y a pas de contradiction car il s'agit de la borne supérieure sur une sous classe  $\Theta_N$  de  $\mathbb{R}^N$ .

Il est temps de préciser ce que contient la classe  $\Theta_N$ . D'une part, si  $\alpha$  est la proportion de coefficients d'amplitudes supérieures ou égales à  $\sigma$ , il vient de

$$\sum_{i=1}^N \min(d_i^2, \sigma^2) = \alpha N \sigma^2 + \sum_{|d_i| < \sigma} d_i^2, \quad (5.9)$$

que  $r_0(\mathbf{d}) \geq \alpha \sigma^2$ . Ainsi, tout signal tel que  $\alpha \geq \zeta(N)$  appartient à  $\Theta_N$ . De ce fait,  $\Theta_N$  peut contenir les représentations en ondelettes de nombreux signaux réguliers par morceaux. En effet, dans le cas d'une représentation en ondelettes, toute singularité crée approximativement le même nombre de coefficients à chaque niveau de décomposition  $j \geq 1$ , tandis que le nombre de coefficients par niveau de décomposition décroît lorsque  $j$  augmente [Mallat, 1999, p. 460].

D'autre part, nous avons la caractérisation suivante (suggérée par un "reviewer" de l'article [Atto, Pastor, and Mercier, 2008a]). Supposons que le signal appartient à un espace de Besov  $B_{p,q}^{s,q}$ ,  $p, q \geq 1$ . D'après [Donoho, 1993b; Donoho and Johnstone, 1996; Krim, Tucker, Mallat, and Donoho, 1999], le risque d'oracle est  $r_0(\mathbf{d}) = O(N^{-2s/(2s+1)})$  tandis que  $\zeta(N) \sim N^{-1/4}$  lorsque  $N$  tend vers l'infini. Cela suggère qu'une grande proportion d'éléments de  $\Theta_N$  appartient à un espace de Besov  $B_{p,q}^s$  tel que  $s < 1/6$ . Les éléments de cette classe sont plutôt non-réguliers puisque le paramètre  $s$  indique, *grosso modo*, le "nombre" de dérivées des éléments dans  $B_{p,q}^s$ . Cela est d'ailleurs confirmé par le fait suivant : le seuil de détection  $\lambda_D(N)$  étant plus petit que les seuils universel et minimax, l'estimée obtenue avec le seuil  $\lambda_D(N)$  sera moins régulière que celle obtenue avec le seuil universel ou minimax. Par la suite, nous allons montrer que la condition  $r_0(\mathbf{d}) \geq \sigma^2 \zeta(N)$  n'est pas très restrictive. En effet, et puisque beaucoup de signaux et d'images naturelles sont plutôt réguliers par morceaux que réguliers,  $\Theta_N$  peut en fait contenir les représentations en ondelettes de nombreux signaux. Cela est confirmé par les résultats expérimentaux donnés à la section suivante : ces résultats montrent que  $r_0(\mathbf{d})$  est plus grand que  $\sigma^2 \zeta(N)$  pour toute image considérée ci-après et la moitié des signaux synthétiques donnés dans WaveLab toolbox de MatLab sont des éléments de  $\Theta_N$ . Un exemple de signal régulier par morceaux appartenant à la classe  $\Theta_N$  pour des valeurs raisonnables de la variance du bruit est le signal 'Blocks' (voir figure 4.1).

Notons aussi que les hypothèses ayant conduit au seuil de détection  $\lambda_D(N)$  peuvent être considérées comme grossières puisque la probabilité de présence des coefficients significatifs est majorée par  $p^* = 1/2$  ( $\lambda_D(N) = \lambda_D(\mathbf{a}, p^*) = \lambda_D(\lambda_u(N), 1/2)$ ). Ce critère a l'avantage de tenir compte des représentations moyennement parcimonieuses. Ainsi, lorsqu'on connaît le degré de régularité du signal à estimer, il est préférable d'ajuster les paramètres du seuil de détection en fonction de la proportion des coefficients significatifs attendus. À titre d'exemple, on sait que les signaux très réguliers admettent des représentations parcimonieuses au sens strict. Ainsi, la proportion des coefficients significatifs est très petite et on peut donc choisir une valeur  $p^*$

très petite devant  $1/2$  : le seuil de détection  $\lambda_D(a, p^*)$ , qui est une fonction décroissante de  $p^*$  devient alors beaucoup plus grand.

Pour finir, remarquons que  $0 \in \mathbb{R}^N$  n'appartient pas à  $\Theta_N$ . Pour estimer  $\mathbf{d} = 0$ , il est évident que plus grand sera le seuil, plus petit sera le risque. Ainsi, les seuils minimax et universels seront naturellement plus adaptés que le seuil  $\lambda_D(N)$  qui est le plus petit des trois. En fait, si  $\mathbf{d} = 0$ , le risque est nul lorsque le seuil est infiniment grand. Ce qui est cohérent avec la Proposition 2 : si seule l'hypothèse nulle peut se réaliser, la probabilité d'occurrence de l'hypothèse alternative est 0. En faisant  $p^* = 0$  dans l'Eq. (5.3), on trouve bien  $\lambda_D(a, 0) = \infty$ . Ce qui illustre le fait qu'une connaissance plus précise de la nature de  $\mathbf{d}$  permet d'obtenir un seuil plus adapté.

## 5.2 Performances des seuils de détection avec la fonction de seuillage doux

La section précédente suggère l'utilisation du seuil  $\lambda_D(N)$  pour une grande classe de signaux, qui est supposée contenir de nombreux signaux réguliers par morceaux. Nous allons vérifier dans cette section que pour une large classe de signaux synthétiques et d'images naturelles, le seuil de détection  $\lambda_D(N)$  conduit à un risque plus faible que les risques associés aux seuils universels et seuils minimax, et cela lorsque l'estimation concerne le seuillage doux. Nous analyserons également l'appartenance des signaux étudiés à la classe  $\Theta_N$ . Plus précisément, nous verrons que le seuil  $\lambda_D(N)$  est systématiquement plus performant que les seuils universel et minimax lorsque la représentation en ondelettes du signal considéré appartient à  $\Theta_N$ . Mais nous verrons aussi que le seuil de détection continue à donner des résultats très satisfaisants même lorsque la représentation en ondelettes du signal considéré n'appartient pas à  $\Theta_N$ . Ce qui confirme le fait mentionné par certains auteurs [Mallat, 1999], [Bruce and Gao, 1996], et qui est que les seuils universels et minimax se révèlent la plupart du temps un peu trop grand en pratique.

### 5.2.1 Résultats expérimentaux sur les signaux

La transformée représentée par la matrice  $\mathcal{W}$  est une TOD basée sur l'ondelette 'sym8' (ondelette Symlet d'ordre 8). Les signaux synthétiques considérés dans cette section sont générés à partir du Wavelet toolbox<sup>1</sup> de Matlab. Comme dans [Donoho and Johnstone, 1994], la taille des échantillons considérés est  $N = 2048$ . Nous posons  $\sigma = 1$  et chaque signal est re-normalisé pour obtenir un Rapport Signal à Bruit (RSB) donné. L'ensemble des RSBs testés est  $\{1, 3, 5, 7\}$ . Le seuillage doux est appliqué aux coefficients de détails des niveaux de décompositions  $j = 1, 2, \dots, J$  où  $J$  est soit 6, soit 10. Les signaux considérés ont différents degrés de parcimonie selon leurs représentations en ondelettes (voir les exemples donnés à la figure 4.1).

Pour tout signal testé, la table 5.3 donne le risque moyen calculé sur 25 réalisations du bruit, lorsque  $\text{RSB} = 1, 3, 5, 7$  et lorsque  $J = 6, 10$ . Nous donnons les résultats pour  $\text{RSB} = 1$  et 7 car les différents résultats obtenus sont assez similaires : pour un RSB donné et pour la plupart des signaux testés, le plus petit risque est obtenu avec le seuil de détection. Notons que concernant les 4 exemples de signaux ('Blocks', 'Doppler', 'Cusp', et 'HypChirps') représentés dans la figure

1. disponible à <http://www-stat.stanford.edu/~wavelab/>



**Tableau 5.3** — Risques  $r_\lambda$  pour le seuillage doux avec les seuils de détection, minimax, et universel. Les signaux écrits en gras sont des éléments de  $\Theta_N$  pour le RSB testé. La valeur  $\zeta(N)$  vaut 0.1035 puisque  $N = 2048$ .

RSB= 1								
$J = 6$					$J = 10$			
Signal	HeaviSine	Bumps	Blocks	Doppler	HeaviSine	Bumps	Blocks	Doppler
$r_{\lambda_D}$	0.0270	0.1758	0.0988	0.0661	0.0423	0.2275	0.1408	0.0967
$r_{\lambda_m}$	0.0227	0.1950	0.1027	0.0677	0.0439	0.2575	0.1558	0.1065
$r_{\lambda_u}$	0.0196	0.3239	0.1386	0.0924	0.0801	0.4461	0.2624	0.1872
Signal	Ramp	Cusp	Sing	HiSine	Ramp	Cusp	Sing	HiSine
$r_{\lambda_D}$	0.0383	0.0231	0.07	0.8421	0.0506	0.0310	0.0794	0.8324
$r_{\lambda_m}$	0.0357	0.0187	0.0743	0.8963	0.0530	0.0301	0.0873	0.8875
$r_{\lambda_u}$	0.0404	0.0157	0.1272	1.0090	0.0917	0.0456	0.1645	1.0059
Signal	LoSine	LinChirp	Piece-Poly	QuadChirp	LoSine	LinChirp	Piece-Poly	QuadChirp
$r_{\lambda_D}$	0.7009	0.7320	0.0727	0.5950	0.6930	0.7273	0.1213	0.5977
$r_{\lambda_m}$	0.7769	0.7877	0.0730	0.6483	0.7692	0.7830	0.1332	0.6537
$r_{\lambda_u}$	0.9941	0.9455	0.09	0.8296	0.9879	0.9472	0.2241	0.8530
Signal	MishMash	Werner	Leopold	Piece-Reg	MishMash	Werner	Leopold	Piece-Reg
$r_{\lambda_D}$	0.7462	0.7247	0.0594	0.0672	0.7465	0.7466	0.0567	0.1158
$r_{\lambda_m}$	0.7912	0.7670	0.0620	0.0672	0.7963	0.7950	0.06	0.1286
$r_{\lambda_u}$	0.9128	0.8758	0.1024	0.0833	0.9392	0.9310	0.1029	0.2262
Signal	Riemann	HypChirps	sineoverx	Chirps	Riemann	HypChirps	sineoverx	Chirps
$r_{\lambda_D}$	0.2744	0.3535	0.0994	0.6976	0.3274	0.3525	0.1252	0.6956
$r_{\lambda_m}$	0.2883	0.3978	0.1077	0.7524	0.3558	0.3978	0.1401	0.7546
$r_{\lambda_u}$	0.3477	0.6320	0.1505	0.9290	0.5048	0.6492	0.2291	0.9428

RSB= 7								
$J = 6$					$J = 10$			
Signal	HeaviSine	Bumps	Blocks	Doppler	HeaviSine	Bumps	Blocks	Doppler
$r_{\lambda_D}$	0.0693	0.4593	0.3644	0.1519	0.1140	0.5373	0.4420	0.1944
$r_{\lambda_m}$	0.0716	0.5390	0.4236	0.1716	0.1291	0.6354	0.5208	0.2268
$r_{\lambda_u}$	0.1032	1.0882	0.8331	0.3284	0.2475	1.3096	1.0692	0.4715
Signal	Ramp	Cusp	Sing	HiSine	Ramp	Cusp	Sing	HiSine
$r_{\lambda_D}$	0.0726	0.0357	0.1185	2.8721	0.1084	0.0629	0.1435	2.8724
$r_{\lambda_m}$	0.0773	0.0334	0.1339	3.4361	0.1237	0.0685	0.1674	3.4382
$r_{\lambda_u}$	0.1353	0.0434	0.2659	7.4357	0.2539	0.1357	0.3567	7.4345
Signal	LoSine	LinChirp	Piece-Poly	QuadChirp	LoSine	LinChirp	Piece-Poly	QuadChirp
$r_{\lambda_D}$	2.2457	2.4482	0.3114	1.8587	2.2193	2.4580	0.3777	1.8845
$r_{\lambda_m}$	2.5826	2.9456	0.3592	2.2344	2.5575	2.9591	0.4446	2.2680
$r_{\lambda_u}$	4.5711	6.4238	0.6736	4.8663	4.5570	6.4559	0.9005	4.9504
Signal	MishMash	Werner	Leopold	Piece-Reg	MishMash	Werner	Leopold	Piece-Reg
$r_{\lambda_D}$	3.6411	3.7349	0.1092	0.2457	3.6485	3.7975	0.1243	0.3204
$r_{\lambda_m}$	4.3826	4.4775	0.1220	0.2803	4.3948	4.5593	0.1422	0.3739
$r_{\lambda_u}$	9.4769	9.4820	0.2326	0.5141	9.5406	9.6995	0.2892	0.7471
Signal	Riemann	HypChirps	sineoverx	Chirps	Riemann	HypChirps	sineoverx	Chirps
$r_{\lambda_D}$	2.6154	0.9497	0.3157	3.1823	2.6693	0.9525	0.3574	3.1960
$r_{\lambda_m}$	3.0237	1.1304	0.3756	3.7977	3.0936	1.1374	0.4289	3.8171
$r_{\lambda_u}$	5.1670	2.3892	0.8042	7.8705	5.3815	2.4288	0.9404	7.9318



4.1, seul le signal 'Cusp' constitue une exception en ce sens où, en fonction du RSB et du niveau de décomposition  $J$ , les performances des seuils minimax et universels peuvent dépasser celles du seuil de détection pour le risque de l'estimation de ce signal.

Notons que pour les valeurs du RSB = 1, 3, 5, 7, environ la moitié des signaux considérés appartiennent à  $\Theta_N$ . Les signaux appartenant à  $\Theta_N$  sont écrits en gras dans la table 5.3. En ce qui concerne les 4 signaux donnés à la figure 4.1, 'Blocks' appartient à  $\Theta_N$  pour RSB = 5, 7, 'Doppler' et 'Cusp' n'appartiennent pas à  $\Theta_N$ , et 'HyChirps' appartient à  $\Theta_N$  pour tous les RSB = 1, 3, 5, 7.

Les tests résumés dans la table 5.3 montrent donc que le seuil de détection donne des meilleures performances lorsque le signal appartient à  $\Theta_N$  pour tous les RSB = 1, 3, 5, 7. Pour les signaux n'appartenant pas à  $\Theta_N$ , le seuil de détection continue à être plus performant que les seuils universel et minimax, sauf pour le signal 'Cusp' où le seuil minimax est le plus performant, car il conduit au plus petit risque. Ce résultat n'est pas très surprenant car la représentation en ondelettes du signal 'Cusp' est très parcimonieuse, comme on peut le voir sur la figure 4.1.

### 5.2.2 Résultats expérimentaux sur les images

Nous considérons les images 'House' et 'Peppers' de dimensions  $256 \times 256$  ainsi que les images 'Barbara', 'Lena', 'Finger', et 'Boat' de dimensions  $512 \times 512$  (voir figure 4.2) utilisées de manière standard pour illustrer le comportement d'algorithmes de traitement d'images. Ces images sont décomposées par la TOD standard de dimension 2. Comme dans la section 5.2.1, nous utilisons l'ondelette 'sym8' pour la décomposition, avec les niveaux de décompositions  $j = 1, 2, \dots, J$  où  $J$  est choisi égal à 4.

On peut constater, à partir de la figure 4.2 que les représentations en ondelettes des images naturelles considérées ne sont pas strictement parcimonieuses. Cette remarque est valable pour la plupart des images naturelles dès qu'elles sont texturées ou représentent des objets (présence de contours). En effet, les textures et les contours des images sont caractérisées par des changements brusques de niveau de gris et traduisent des singularités. Ainsi, on peut s'attendre à ce que les représentations des images naturelles (comme celles considérées pour les tests) appartiennent à  $\Theta_N$  et donc, que le seuil de détection soit plus performant que les seuils universels et minimax.

On vérifie (voir la table 5.4) que  $r_0(\mathbf{d}) \geq \sigma^2 \zeta(N)$  de sorte que  $\mathbf{d} \in \Theta_N$  pour toutes les images testées et tous les écart-type de bruit  $\sigma = 9, 18, 27, 36$  considérés. Plus précisément, on peut remarquer que  $r_0(\mathbf{d})$  est généralement beaucoup plus grand que  $\sigma^2 \zeta(N)$  pour une bonne partie des images testées. Nous rappelons ici que pour chaque image,  $\mathbf{d}$  représente ici le vecteur formé par l'ensemble des coefficients de la TOD bi-dimensionnelle.

La table 5.5 présente les risques obtenus pour les seuils de détection, minimax et universel, lorsque l'estimation par seuillage doux est appliquée aux coefficients de détails des niveaux de décomposition  $j = 1, 2, \dots, J = 4$ . L'estimation porte sur des images corrompues par du bruit additif blanc Gaussien d'écart-type  $\sigma = 9, 18, 27$ , et 36. Les risques donnés dans la table 5.5 sont les risques moyens obtenus à partir de 10 essais sur des réalisations différentes du bruit additif blanc Gaussien. Comme on peut le constater, ces risques sont plus faibles lorsque le seuil de détection est utilisé, ce qui confirme le fait que ce seuil est plus adapté que les seuils universel et minimax pour l'estimation par seuillage doux des éléments de  $\Theta_N$ .

**Tableau 5.4** — Valeurs  $r_0(\mathbf{d})$  pour les images testées. Chaque vecteur  $\mathbf{d}$  correspond aux coefficients de la TOD d'une image testée. La TOD est basée sur l'ondelette 'sym8', avec  $J = 4$ .

$N$	$256 \times 256$		$512 \times 512$			
Images	'House'	'Peppers'	'Barbara'	'Lena'	'Finger'	'Boat'
$\sigma = 9$						
$\sigma^2 \zeta(N)$	3.5412	3.5412	2.5070	2.5070	2.5070	2.5070
$r_0(\mathbf{d})$	17.2998	24.4012	25.0542	17.9539	36.0245	26.5791
$\sigma = 18$						
$\sigma^2 \zeta(N)$	14.1647	14.1647	10.0280	10.0280	10.0280	10.0280
$r_0(\mathbf{d})$	37.9459	56.1343	60.4723	36.4248	87.0434	55.7944
$\sigma = 27$						
$\sigma^2 \zeta(N)$	31.8705	31.8705	22.5630	22.5630	22.5630	22.5630
$r_0(\mathbf{d})$	58.3006	89.8592	98.1831	54.9842	139.9741	84.0878
$\sigma = 36$						
$\sigma^2 \zeta(N)$	56.6587	56.6587	40.1120	40.1120	40.1120	40.1120
$r_0(\mathbf{d})$	78.6185	124.2232	136.0478	73.5691	193.8070	111.5801

**Tableau 5.5** — Risques  $r_\lambda$  obtenus avec les seuils de détection, minimax, et universel pour l'estimation par seuillage doux (TOD - 'sym8' -  $J = 4$ ).

$N$	$256 \times 256$		$512 \times 512$			
Images	'House'	'Peppers'	'Barbara'	'Lena'	'Finger'	'Boat'
$\sigma = 9$						
$r_{\lambda_D(N)}$	46.5527	71.4106	89.4982	45.8152	119.5679	74.5194
$r_{\lambda_m(N)}$	60.7369	97.2327	130.4604	62.7722	173.3831	103.3656
$r_{\lambda_u(N)}$	84.3010	141.5590	180.4511	82.9846	240.7068	137.0883
$\sigma = 18$						
$r_{\lambda_D(N)}$	89.3033	150.7902	189.6661	87.2403	255.6930	143.1684
$r_{\lambda_m(N)}$	114.4277	199.3795	257.5655	116.3257	358.1196	189.5936
$r_{\lambda_u(N)}$	153.6769	279.7856	325.7944	149.0234	481.0768	238.7614
$\sigma = 27$						
$r_{\lambda_D(N)}$	128.7033	225.4500	269.5180	124.0678	384.6643	200.0341
$r_{\lambda_m(N)}$	160.9144	294.0141	343.2329	160.7170	529.2084	255.7203
$r_{\lambda_u(N)}$	211.2654	403.9108	404.9466	199.0735	699.8262	311.7397
$\sigma = 36$						
$r_{\lambda_D(N)}$	164.7825	297.3454	328.7292	156.6675	507.7654	247.9232
$r_{\lambda_m(N)}$	202.1105	380.8619	399.0873	197.1790	689.4484	308.4268
$r_{\lambda_u(N)}$	262.7180	508.8113	452.3670	237.0470	898.5787	363.9042

### 5.3 Débruitage en utilisant une transformée en ondelettes redondante

La TOD invariante par translation [Coifman and Donoho, 1995], plus connue sous le nom de SWT (pour Stationary Wavelet Transform), est utilisée dans cette section pour le débruitage de signaux et d'images. Cette transformée est particulièrement appréciée pour le débruitage puisqu'elle est redondante et sa redondance a pour effet de supprimer le bruit résiduel (dû à la sensibilité de la TOD aux translations) par un effet de moyennage. Les tests sont effectués avec l'ondelette 'sym8' et le seuillage doux concerne les coefficients de détail des niveaux de décompositions  $j = 1, 2, \dots, J$  avec  $J = 6$  pour les signaux et  $J = 4$  pour les images.

#### 5.3.1 Débruitage de signal

Nous considérons le signal 'Doppler' corrompu par du bruit additif blanc Gaussien d'écart-type  $\sigma = 1$  et de RSB = 7. La figure 5.1 montre le signal 'Doppler' bruité, en comparaison avec les signaux 'Doppler' débruités par seuillage doux et utilisation des seuils de détection, minimax et universel. Le signal 'Doppler' original est représenté en pointillé sur chacune des 3 figures représentant les signaux débruités. La figure 5.2 donne un zoom des premiers 50 et des derniers 50 coefficients des versions débruitées données par la figure 5.1. Ces figures montrent que le seuillage doux avec le seuil universel conduit à l'estimée la plus régulière, ce qui est normal car c'est le plus grand des 3 seuils. Cependant, cette régularité n'est pas pour autant synonyme de meilleure estimation car l'estimée obtenue avec le seuil universel est celle qui s'écarte le plus du signal original, comme on peut le voir sur la figure 5.2.

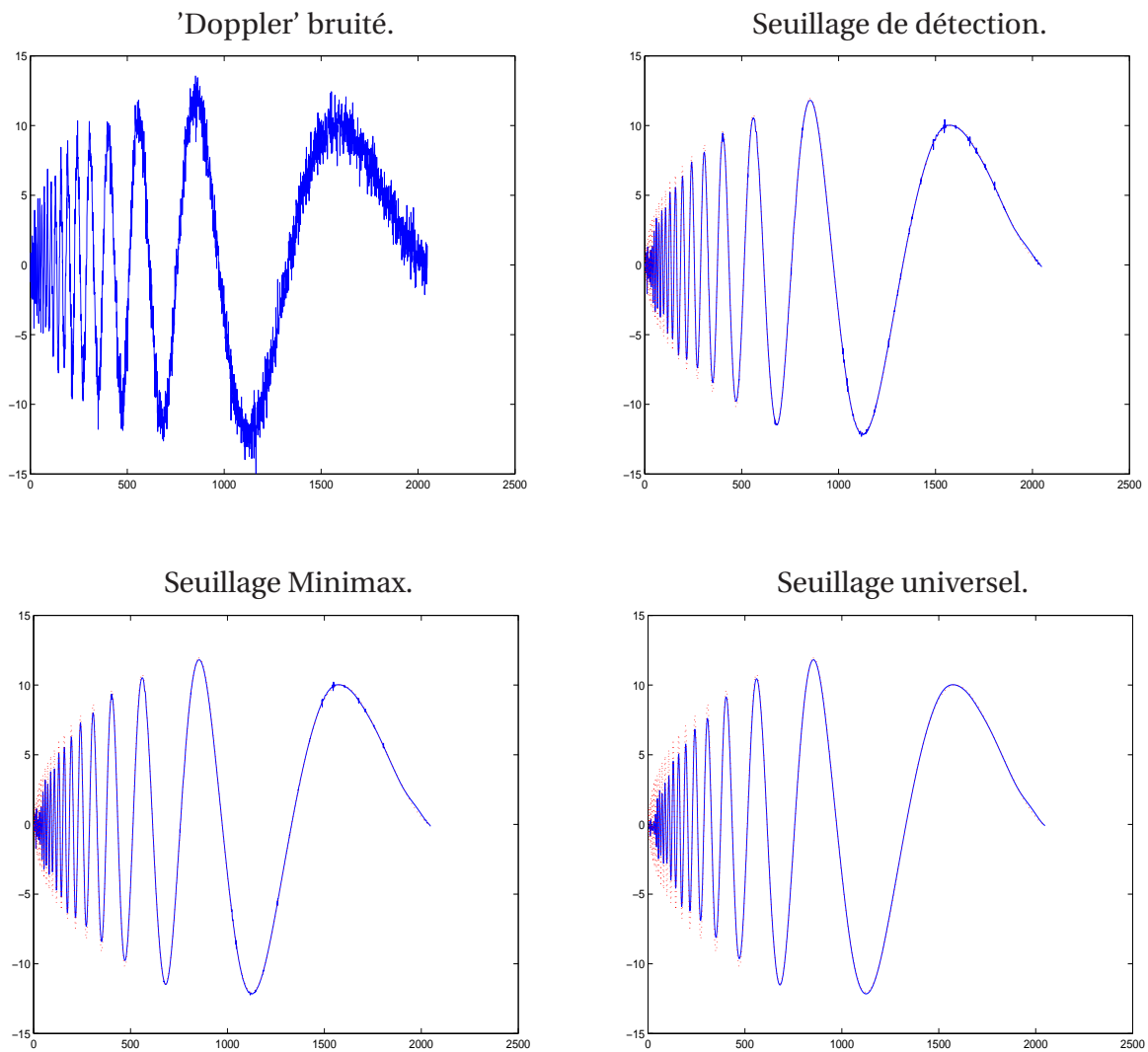
#### 5.3.2 Débruitage d'images

Nous considérons l'image de 'Lena' corrompue par du bruit additif blanc Gaussien. La table 5.6 présente les risques obtenus avec les seuils de détection, minimax et universel, lorsque l'estimation par seuillage doux est utilisée pour débruiter cette image. La table 5.6 donne les risques moyens obtenus (sur 10 tests) et la table 5.7 donne les PSNRs correspondants, pour les valeurs de  $\sigma$  considérées. L'expression du PSNR (pour Peak Signal-to-Noise Ratio) est donné, en deciBel (dB), par

$$\text{PSNR}(\lambda) = 10 \log_{10} (255^2 / r_\lambda). \quad (5.10)$$

où  $r_\lambda$  désigne le risque de l'estimation par seuillage doux avec le seuil  $\lambda$ . Le PSNR est un critère couramment utilisé pour juger de la qualité d'une image débruitée ou compressée. On constate à partir des résultats de la table 5.7 qu'avec le seuil de détection, le gain est de l'ordre de 1 à 2 dBs, comparativement aux PSNRs obtenus avec le seuil minimax, qui lui même est plus performant d'environ 1 dB que le seuil universel.

Un exemple de débruitage de l'image 'Lena' est donné par la figure 5.3. L'écart-type du bruit est  $\sigma = 25$ . Cette figure montre que l'image débruitée avec le seuil de détection est de meilleure qualité (les contours et les textures sont mieux restitués, l'image est plus nette) que celles obtenues avec les seuils universels et minimax.

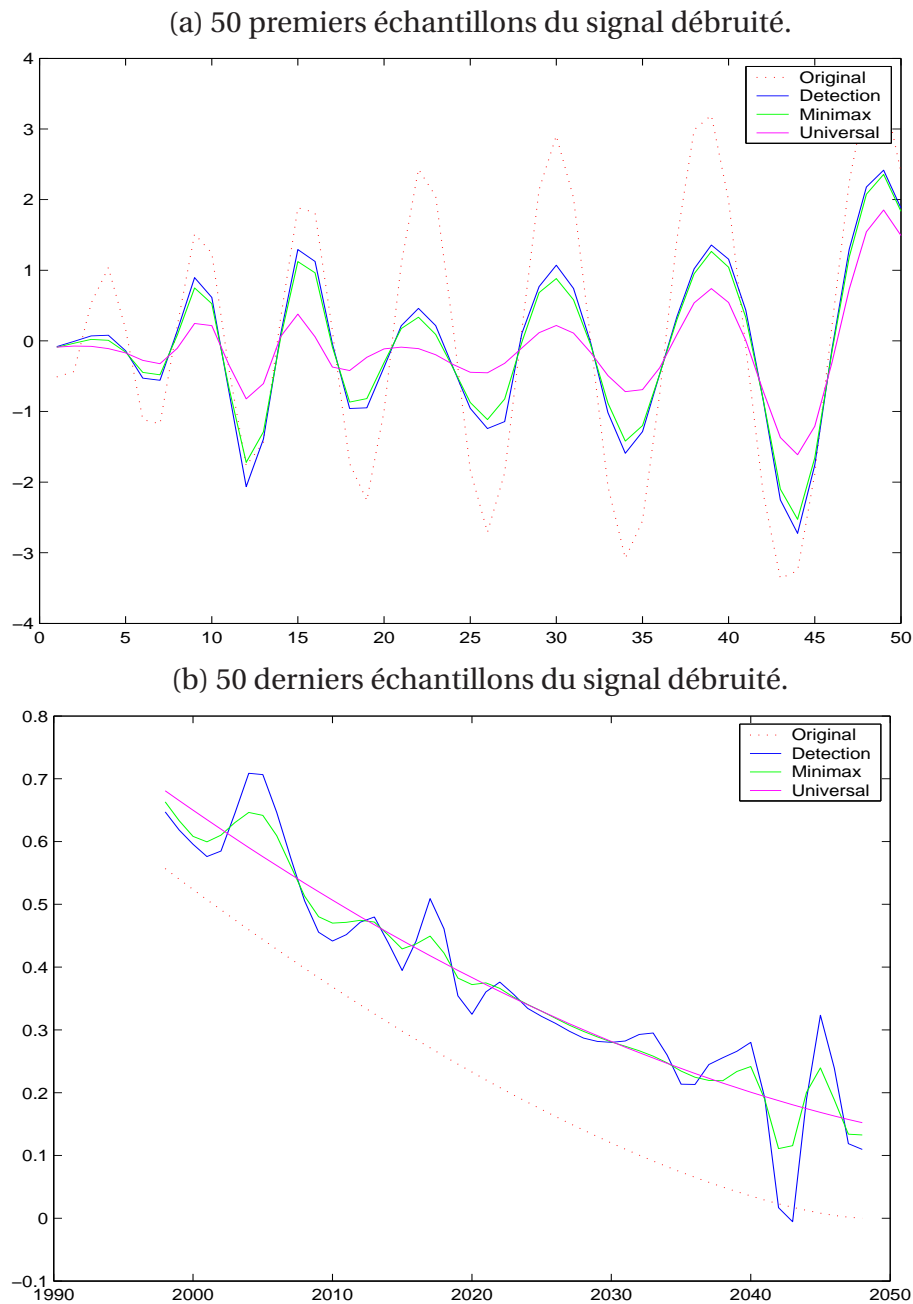


**Figure 5.1** — Débruitage du signal 'Doppler' par SWT 'sym8' avec  $J = 6$  et par seuillage doux.

**Tableau 5.6** — Risques obtenus avec les seuils de détection, minimax, et universel pour le débruitage par seuillage doux des coefficients de détails de l'image bruitée de 'Lena' (SWT - 'sym8' -  $J = 4$ ).

$\sigma$	9	18	27	36
$r_{\lambda_u(N)}$	76.0	141.6	192.1	231.3
$r_{\lambda_m(N)}$	56.2	108.7	152.8	189.3
$r_{\lambda_D(N)}$	40.0	79.4	115.1	146.5

**Remarque 4** La stratégie de seuillage adopté ici est globale en ce sens où le seuillage ne tient pas compte de la structure de la transformée en ondelettes. Cependant, on sait, [Donoho, 1993a], que l'estimation est plus performante lorsque les seuils sont calculés en fonction de la répartition des coefficients par niveau de décomposition, plutôt qu'en appliquant un seuil global calculé sur tous les coefficients d'ondelettes. Aussi, dans [Pastor and Atto, 2008], le fait d'utiliser un seuillage doux avec les seuils  $\lambda_D^j(N)$  tels que  $\lambda_D^1(N) = \lambda_D(N)$  et  $\lambda_D^{j+1}(N) = \lambda_D^j(N)/\sqrt{2}$  (voir aussi [Donoho,



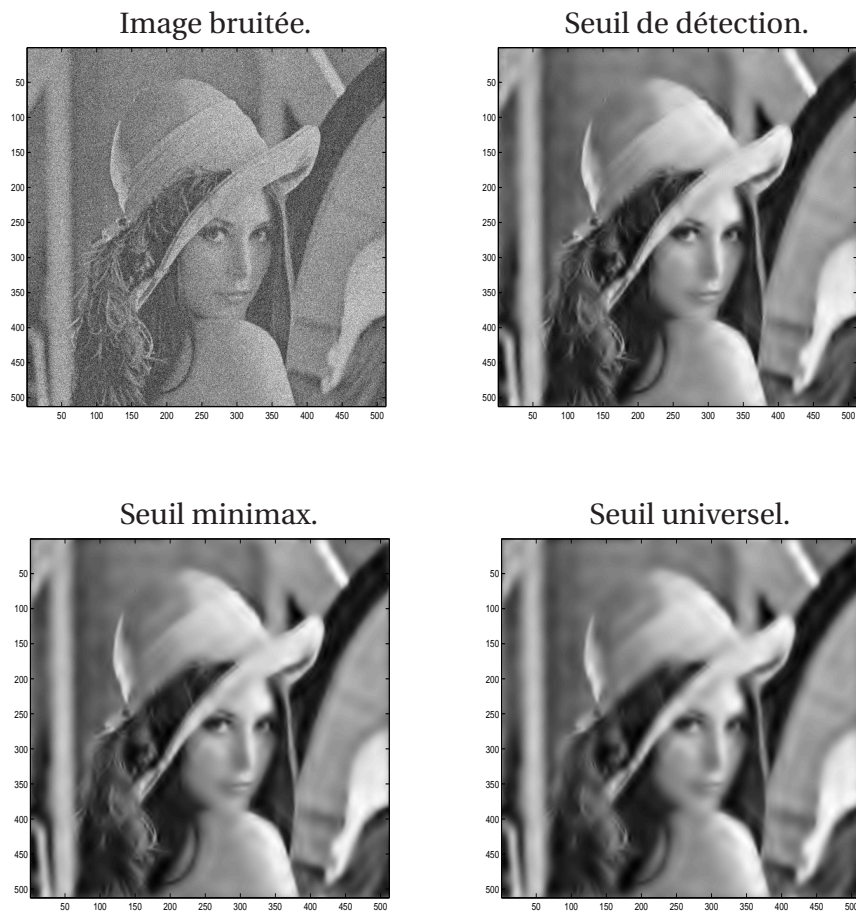
**Figure 5.2** — Zooms des (a) 50 premiers échantillons du signal (b) 50 derniers échantillons - du signal 'Doppler' et de ses versions débruitées par les seuils de détection, minimax et universel.

*1993a)* conduit à des résultats d'environ 2 dBs plus satisfaisants que ceux obtenus avec le seuil  $\lambda_D(N)$ .

La section suivante propose la synthèse de seuils de détection plus adaptés à la nature de la décomposition en ondelettes d'un signal en ce sens où les bornes  $a$  et  $p^*$  sont calculés à partir de la parcimonie locale (par niveau de décomposition) des coefficients d'ondelettes.

**Tableau 5.7** — PSNRs obtenus avec les seuils de détection, minimax, et universel pour le débruitage par seuillage doux des coefficients de détails de l'image bruitée de 'Lena' (SWT - 'sym8' -  $J = 4$ ).

PSNR initial	29.0	23.0	19.5	17.0
PSNR[ $\lambda_u(N)$ ]	29.3	26.6	25.3	24.5
PSNR[ $\lambda_m(N)$ ]	30.6	27.8	26.3	25.4
PSNR[ $\lambda_D(N)$ ]	32.1	29.1	27.5	26.5



**Figure 5.3** — Image bruitée et images débruitées de 'Lena', par seuillage doux avec les seuils de détection, minimax et universel (SWT - 'sym8' -  $J = 4$ ). L'écart-type du bruit est  $\sigma = 36$ .

### 5.3.3 Amélioration des performances

Le seuil  $\lambda_D(N)$  est un seuil de détection calculé avec très peu de contraintes sur le degré de parcimonie de la représentation en ondelettes du signal. En effet, d'une part, l'amplitude minimale des coefficients dits significatifs pour le signal est supposée ne pas être plus grande que le niveau du maximum du bruit :  $a = \lambda_u(N)$ . D'autre part, la probabilité de présence des coefficients du signal peut atteindre  $p^* = 1/2$ , ce qui revient à dire que l'on considère comme représentation la moins parcimonieuse (mais parcimonieuse tout de même), celle dont la proportion des grands coefficients du signal est d'environ  $1/2$ .

Cependant, il faut noter ceci. Tout d'abord, on peut choisir l'amplitude minimale significative  $a$  et la borne  $p^*$  de la probabilité de présence du signal en fonction de l'appartenance de ce signal à un espace de Besov donné (la parcimonie de la représentation du signal est en effet liée à la régularité de ce signal). De plus, on peut également choisir ces bornes en tenant compte de la spécificité de la représentation en ondelettes et selon le niveau de décomposition  $j$  considéré (voir remarque 4).

Pour ne pas introduire de connaissance *a priori* sur le signal que nous voulons estimer, nous considérons ci-dessous, une adaptation de ces bornes en fonction de la spécificité de la TOD. En effet, compte tenu de [Mallat, 1999, Theorem 9.7, p. 389], il est légitime de considérer que pour une TOD de signaux réguliers ou réguliers par morceaux, l'amplitude minimale significative est telle que  $a_j = a_{j-1} / \sqrt{2}$ , où  $j$  est le niveau de décomposition. On sait également que dans le cas d'un signal régulier par morceaux, la proportion des coefficients significatifs est moins importante au niveau de décomposition  $j$  qu'au niveau de décomposition  $j+1$ . Cela est justifié par le fait qu'une singularité crée approximativement le même nombre de coefficients significatifs aux niveaux de décompositions  $j$  et  $j+1$ , tandis qu'il y'a 2 fois moins de coefficients au niveau de décomposition  $j+1$  qu'au niveau de décomposition  $j$ . Ainsi,  $p_j^*$  doit être une fonction croissante de  $j$ . À noter qu'il faudrait arrêter le seuillage à un niveau  $J$  pour lequel  $p_J^*$  est de l'ordre de  $1/2$ . Nous posons donc que  $p_J^* = 1/2$ . Nous proposons l'utilisation de séquences géométriquement ou exponentiellement croissante de  $j$  pour le choix de  $p_j^*$ , puisque  $p_1^*$  doit être une valeur très petite (il y'a en général très peu d'information significative sur le signal à  $j=1$ ) et la présence d'information significative augmente très vite avec le niveau de décomposition. Un choix pourrait donc être  $p_{j+1}^* = (p_j^*)^{1/\nu}$ , où  $\nu > 1$  est une valeur à choisir en fonction de la nature du signal analysé.

Ces hypothèses, raisonnables compte tenu de la nature de la décomposition en ondelettes, conduisent aux seuils  $\lambda_D(a_j, p_j^*)$  qui sont plus adaptés pour la distinction entre les coefficients du signal et ceux du bruit, lorsque la représentation est une TOD. En résumé, nous considérons donc

$$a_j = a_1 / \sqrt{2^{j-1}}, \quad (5.11)$$

et

$$p_j^* = 1/2^{\nu^{j-1}}. \quad (5.12)$$

Dans le chapitre suivant, nous illustrerons l'efficacité de telles approches avec une fonction d'atténuation plus appropriée que la fonction de seuillage doux. Nous considérerons alors  $a_1 = \lambda_u(N)$ , de sorte que

$$a_j = \sigma \sqrt{\ln N} / 2^{\frac{j}{2}-1}. \quad (5.13)$$



## Estimation non-paramétrique par atténuation sigmoïdale

### 6.1 Contexte de la méthode

Les fonctions SSBS (pour Smooth Sigmoid-Based Shrinkage) sont des fonctions régulières qui opèrent une atténuation sigmoïdale sur les données. Ces fonctions ont été présentées dans [Atto, Pastor, and Mercier, 2008c] dans le but de corriger les limitations des fonctions de seuillages classiques telles que celles représentées à la figure 4.4. Notons qu'en pratique, les observations suivantes ont été signalées dans le cas de l'estimation par seuillage dans le domaine des ondelettes.

Les estimées obtenues par seuillages dur et doux admettent respectivement une grande variance et un biais important dus essentiellement au seuillage [Bruce and Gao, 1996, page 5]. Certaines fonctions comme les fonctions NNG et SCAD données à la figure 4.4 réalisent un compromis entre le seuillage doux et le seuillage dur. Cependant, toutes ces fonctions sont basées sur le seuillage, ce qui introduit un traitement singulier en ce sens où, de part et d'autre du seuil, les données sont traitées de manière différente.

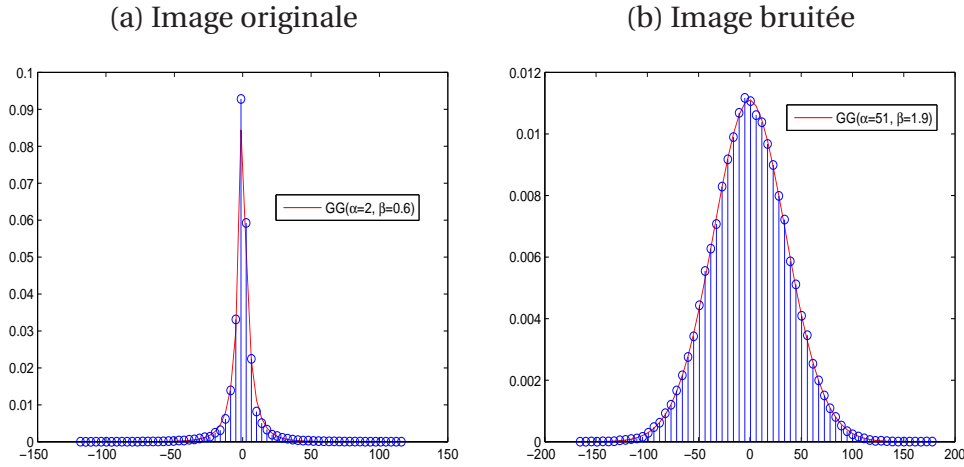
Notons que le principe du seuillage assume que la représentation du signal est parcimonieuse au sens strict. Si cela est souvent le cas pour des signaux très réguliers, il n'en est pas de même pour les signaux réguliers par morceaux comme les images naturelles (voir figure 4.2) ou des signaux encore plus complexes comme le signal de parole. Le fait que ces derniers signaux admettent une représentation en ondelettes parcimonieuse au sens large ne justifie pas l'adéquation du seuillage pour l'estimation de ces signaux. Plus précisément, ce qui suit illustre l'effet du seuillage sur la distribution des coefficients d'ondelettes d'une image naturelle.

La figure 6.1 montre les histogrammes des coefficients d'ondelettes des détails horizontaux de l'image de 'Lena' et de l'image bruitée de 'Lena' (premier niveau de décomposition). Ces histogrammes sont superposés avec des distributions Gaussiennes Généralisées (GG). On rappelle qu'une distribution GG d'échelle  $\alpha$  et de forme  $\beta$  est, d'après l'Eq. (3.12) :

$$f_{\alpha,\beta}(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta}.$$

Les distributions GG décrivent parfois plus ou moins bien la nature des coefficients d'ondelettes de certaines images naturelles, comme on peut le voir sur la figure 6.1 (a) (voir également [Sendur and Selesnick, 2002; Do and Vetterli, 2002]).





**Figure 6.1** — Histogrammes des coefficients de détails verticaux du premier niveau de décomposition d’une image ‘Lena’ et de sa version bruitée par du bruit additif blanc Gaussien d’écart-type  $\sigma = 35$ . Les histogrammes sont superposés avec des distributions GG dont les paramètres sont donnés sur la figure. La SWT basée sur l’ondelette de Haar a été utilisée pour la décomposition.

Le seuillage des coefficients de détails de l’image bruitée (coefficients représentés par l’histogramme de la figure 6.1 (b)) par les fonctions de seuillage classiques (celles représentées à la figure 4.4) donne le résultat représenté à la figure 6.2. Dans cette figure, ainsi que dans le reste du document,  $G_h$  désigne le traitement (seuillage, atténuation) effectué par la fonction  $G$  avec l’ensemble des paramètres donné par  $h$ . Ainsi,  $\text{SCAD}_\lambda$  signifie que l’estimation est réalisée à l’aide de la fonction de seuillage SCAD, calibrée avec le seuil  $\lambda$ . Notons que le seuillage utilisé correspond à l’écart-type du bruit ( $\lambda = \sigma$ ). Nous avons choisi cette valeur pour bien mettre en évidence l’effet du seuillage. En fait, les seuils universels, minimax, et de détection étant tous plus grands que  $\sigma$ , le résultat est encore plus mauvais : les distributions obtenues deviennent très singulières en ce sens où elles sont caractérisées par un très fort pic en 0 et quelques rares coefficients éparpillés de part et d’autre de ce pic. Ainsi, il n’est pas possible d’obtenir des distributions qui puissent se rapprocher de la distribution des coefficients de l’image originale, représentée par la figure 6.1 (a).

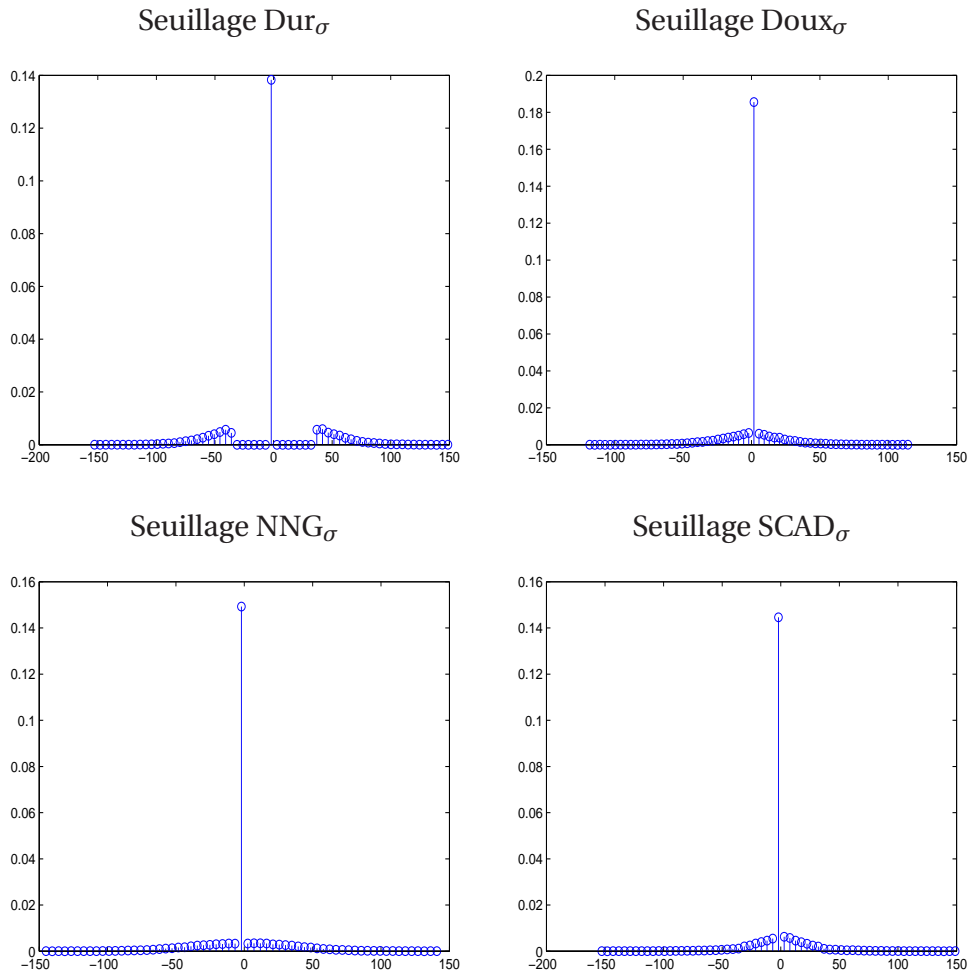
Nous traitons dans la section suivante, le problème du choix de fonctions d’atténuation de manière à pouvoir réduire le bruit sans trop affecter la distribution du signal. Pour réaliser de tels traitements, et compte tenu de la parcimonie des représentations en ondelettes (voir 4.1 et 4.2), nous considérons qu’une fonction adaptée à l’estimation des coefficients du signal est une fonction qui possède les propriétés de :

**(P1) Régularité** : l’atténuation d’amplitude doit être régulière de manière à introduire moins de variabilité entre coefficients d’amplitudes voisines.

**(P2) Atténuation différentielle** : une grande (resp. faible) atténuation pour les petits (resp. grands) coefficients est requise puisque les petits (resp. grands) coefficients contiennent moins (resp. plus) d’information sur le signal.

D’autre part, la troisième propriété suivante peut être souhaitée dans certaines circonstances,

**(P3) Annulation de l’atténuation** : pour éviter un lissage trop important de l’estimée, l’at-



**Figure 6.2** — Histogrammes obtenus après seuillage des coefficients de l’image bruitée ‘Lena’. Les coefficients seuillés sont ceux représentés par l’histogramme de la figure 6.1 (b). Le seuil utilisé est  $\lambda = \sigma$ , l’écart-type du bruit.

ténuation doit décroître vers zéro lorsque l’amplitude des coefficients tend vers l’infini.

## 6.2 Définition, propriétés des fonctions SSBS

Les fonctions SSBS sont définies pour tout  $x$  réel par [Atto, Pastor, and Mercier, 2008c] :

$$\delta_{t,\tau,\lambda}(x) = \frac{\text{sgn}(x)(|x| - t)_+}{1 + e^{-\tau(|x| - \lambda)}}, \quad (6.1)$$

pour  $(t, \tau, \lambda) \in \mathbb{R}_+ \times \mathbb{R}_+^* \times \mathbb{R}_+$ , avec  $\text{sgn}(x) = 1$  (resp.  $-1$ ) si  $x \geq 0$  (resp.  $x < 0$ ), et  $(x)_+ = x$  (resp.  $0$ ) si  $x \geq 0$  (resp.  $x < 0$ ). Chaque fonction  $\delta_{t,\tau,\lambda}$  est le produit d’une fonction de seuillage doux par une fonction de type sigmoïd. C’est ce qui justifie le nom de SSBS (Smooth Sigmoid-Based Shrinkage) attribué à ces fonctions.

Il est facile de vérifier que ces fonctions satisfont aux conditions (P1) et (P2). D’autre part, les fonctions de seuillage dur et doux peuvent être vue comme étant des fonctions  $\delta_{t,\tau,\lambda}$  pour

des valeurs dégénérées du paramètre  $\tau$ . En effet, pour  $t = \lambda > 0$  fixé, la fonction  $\delta_{t,\tau,t}(x)$  tend vers la fonction de seuillage doux  $\text{sgn}(x)(|x| - t)_+$  lorsque  $\tau$  tend vers  $+\infty$ . D'autre part, si l'on pose  $t = 0$  et si l'on fait tendre  $\tau$  vers  $+\infty$ , alors  $\delta_{0,\tau,\lambda}(x)$  tend vers  $\delta_{0,\infty,\lambda}(x)$ , qui est la fonction de seuillage dur donnée par :

$$\delta_{0,\infty,\lambda}(x) = \begin{cases} x & \text{if } |x| > \lambda, \\ \pm \lambda/2 & \text{if } x = \pm \lambda, \end{cases} \quad (6.2)$$

où  $\Delta$  est la fonction indicatrice d'intervalle  $\Delta \subset \mathbb{R}$  définie par :  $\Delta(x) = 1$  si  $x \in \Delta$  ;  $\Delta(x) = 0$  si  $x \in \mathbb{R} \setminus \Delta$ . Notons que  $\delta_{0,\infty,\lambda}$  réduit un coefficient d'amplitude  $\lambda$  à la moitié de sa valeur, et de ce fait, minimise la variation locale autour de  $\lambda$ , puisque  $\lim_{x \rightarrow \lambda^+} \delta_{0,\infty,\lambda}(x) - 2\delta_{0,\infty,\lambda}(\lambda) + \lim_{x \rightarrow \lambda^-} \delta_{0,\infty,\lambda}(x) = 0$ .

Pour des valeurs finies fixées de  $t$ ,  $\tau$  et  $\lambda$ ,  $\delta_{t,\tau,\lambda}(x)$  se comporte comme le seuillage doux  $\text{sgn}(x)(|x| - t)_+$  lorsque  $x \rightarrow \infty$ . Ainsi, le paramètre  $t$  impose une atténuation lorsque  $x$  est grand (voir figure 6.3). Une telle atténuation peut ne pas être désirée lorsque  $t$  est grand car cette atténuation peut générer un lissage trop important de l'estimée. C'est le cas lorsque le seuillage doux est effectué avec le seuil universel ou le seuil minimax, comme on peut le voir sur la figure 5.3.

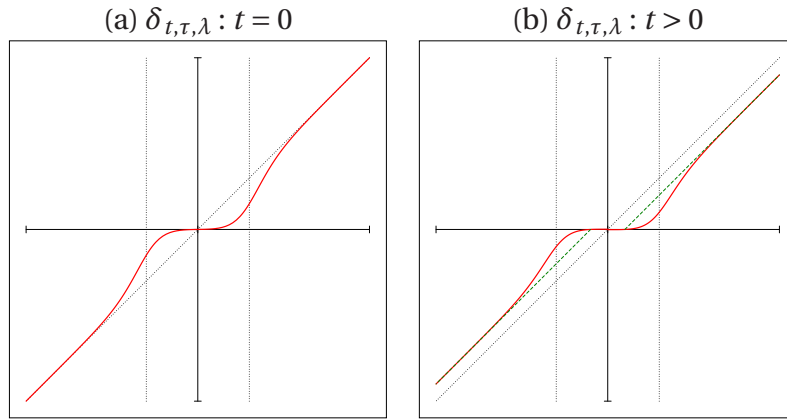


Figure 6.3 — Graphes des fonctions  $\delta_{t,\tau,\lambda}$  pour  $t = 0$  et  $t \neq 0$ .

Par conséquent, pour éviter de telles situations, nous considérons par la suite, le cas  $t = 0$  :  $\delta_{\tau,\lambda} = \delta_{0,\tau,\lambda}$ , ce qui conduit à

$$\delta_{\tau,\lambda}(x) = \frac{x}{1 + e^{-\tau(|x| - \lambda)}}. \quad (6.3)$$

Les fonctions  $\delta_{\tau,\lambda}$  données par (6.3) satisfont à la fois aux propriétés (P1), (P2), et garantissent également la condition (P3).

Le paramètre  $\lambda$  joue un rôle analogue à celui d'un seuil puisque  $\delta_{\infty,\lambda}$  est une fonction de seuillage dur de hauteur  $\lambda$ . On vérifie d'autre part que les points  $A = (\lambda, \lambda/2)$ ,  $O = (0, 0)$  et  $A' = (-\lambda, -\lambda/2)$ , en coordonnées cartésiennes, appartiennent à la courbe de la fonction  $\delta_{\tau,\lambda}$  pour tout  $\tau > 0$ . En effet, on a  $\delta_{\tau,\lambda}(\pm\lambda) = \pm\lambda/2$  et  $\delta_{\tau,\lambda}(0) = 0$  pour tout  $\tau > 0$ , d'après l'Eq. (6.3). Il s'ensuit que  $\tau$  paramétrise la courbure de l'arc  $\overline{A'O A}$  porté par la courbe de la fonction SSBS dans l'intervalle  $] -\lambda, \lambda[$ . Cette courbure est directement liée au degré d'atténuation que nous voulons imposer aux coefficients d'ondelettes.

Considérons alors l'arc  $\widehat{OA}$  du graphe de la figure 6.4 (figure qui représente une fonction SSBS tracée dans le demi-plan positif). Soit  $C$ , l'intersection entre l'axe des abscisses

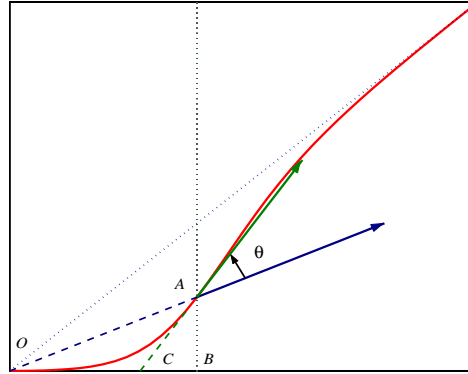


Figure 6.4 — Graphe de la fonction  $\delta_{\tau, \lambda}$  dans le demi-plan positif.

et la tangente à la courbe de la fonction SSBS au point  $A$ . L'équation de cette tangente est  $y = 0.25(2 + \tau\lambda)(x - \lambda) + 0.5\lambda$ . Les coordonnées du point  $C$  sont  $C = (\tau\lambda^2/(2 + \tau\lambda), 0)$ . On peut facilement contrôler la courbure de l'arc  $\widehat{OA}$  grâce à l'angle noté  $\theta$ , entre le vecteur fixe  $\overrightarrow{OA}$  et le vecteur  $\overrightarrow{CA}$ , qui est le vecteur porté par la tangente de la courbe de  $\delta_{\tau, \lambda}$  au point  $A$ . Plus  $\theta$  sera grand, plus grande sera l'atténuation imposée aux coefficients d'amplitudes inférieures ou égales à  $\lambda$ . Pour  $\lambda$  fixé, la relation entre l'angle  $\theta$  et le paramètre  $\tau$  est

$$\cos \theta = \frac{\overrightarrow{OA} \cdot \overrightarrow{CA}}{\|\overrightarrow{OA}\| \cdot \|\overrightarrow{CA}\|} = \frac{10 + \tau\lambda}{\sqrt{5(20 + 4\tau\lambda + \tau^2\lambda^2)}}. \quad (6.4)$$

On obtient de l'Eq. (6.4) que  $0 < \theta < \arccos(\sqrt{5}/5)$ ; lorsque  $\theta = \arccos(\sqrt{5}/5)$ , alors  $\tau = +\infty$  et  $\delta_{\tau, \lambda}$  est la fonction de seuillage dur donnée par l'Eq. (6.2). De l'Eq. (6.4), il s'ensuit que  $\tau$  peut être écrit comme une fonction de  $\theta$  et  $\lambda$  par :

$$\tau(\theta, \lambda) = \frac{10}{\lambda} \frac{\sin^2 \theta + 2 \sin \theta \cos \theta}{5 \cos^2 \theta - 1}. \quad (6.5)$$

En pratique, lorsque  $\lambda$  est fixé, l'angle  $\theta$  permet de choisir le degré d'atténuation imposé aux données d'amplitudes appartenant à l'intervalle  $]0, \lambda[$ , ce qui est plus naturel que l'utilisation du paramètre  $\tau$ . Par la suite,  $\delta_{\theta, \lambda} = \delta_{\tau(\theta, \lambda), \lambda}$  désigne une fonction SSBS lorsque  $\tau(\theta, \lambda)$  est donné par l'Eq. (6.5). Compte tenu de cette interprétation des paramètres des fonctions SSBS, il est alors facile de trouver des paramètres adéquats en fonction du résultat recherché. En résumé, la procédure d'estimation consiste en 3 étapes :

1. Fixer le seuil  $\lambda$  et l'angle  $\theta$  de la fonction SSBS, avec  $\lambda > 0$  et  $0 < \theta < \arccos(\sqrt{5}/5)$ . À noter que plus  $\theta$  sera grand, plus forte sera l'atténuation.
2. Calculer la valeur de  $\tau$  à partir de l'Eq. (6.5).
3. Atténuer les données selon la fonction SSBS  $\delta_{\tau, \lambda}$  définie par l'Eq. (6.3).

Dans ce qui suit, les termes “degré d'atténuation” et “seuil” désignent  $\theta$  et  $\lambda$ , respectivement. En outre, la notation  $\delta_{\tau, \lambda}$  sera préférée pour les résultats théoriques ; la notation  $\delta_{\theta, \lambda}$ , introduite

ci-dessus sera utilisée pour les résultats expérimentaux, puisque le choix du degré d'atténuation  $\theta$  est plus naturel que le choix du paramètre  $\tau$ . Quelques graphes des fonctions SSBS sont donnés à la figure 6.5 (à gauche) pour différentes valeurs du degré d'atténuation  $\theta$  (le seuil  $\lambda$  est fixé).

### 6.3 Régularisation avec les fonctions SSBS

Considérons l'estimation des coefficients du signal (vecteur  $\mathbf{d}$ ) dans le modèle parcimonieux donné par l'Eq. (4.1). Cette estimation peut être traitée par une approche de régularisation visant à garantir une borne sur la norme de l'EQM. Pour cela, on se donne une fonction de pénalité  $q_\lambda = q_\lambda(\cdot)$  et on cherche le vecteur  $\mathbf{d}$  qui minimise la quantité [Antoniadis and Fan, 2001]

$$\|\mathbf{d} - \mathbf{c}\|_{\ell_2}^2 + 2 \sum_{i=1}^N q_\lambda(|d_i|). \quad (6.6)$$

Le seuillage et la régularisation sont liés dans le sens où une fonction de seuillage correspond à une procédure de régularisation avec une fonction de seuillage spécifique [Antoniadis, 2007]. Cette correspondance est précisée par la proposition suivante :

**Proposition 4** [Antoniadis, 2007, Proposition 3.2].

*Soit  $\delta$  une fonction de seuillage de la variable réelle, antisymétrique, telle que  $0 \leq \delta(x) \leq x$  pour  $x \geq 0$  et  $\delta(x)$  tend vers l'infini lorsque  $x$  tend vers l'infini. Alors il existe une fonction de pénalité continue et positive  $q$ , avec  $q(|x|) \leq q(|y|)$  lorsque  $|x| \leq |y|$ , et telle que  $\delta(z)$  est la solution unique du problème de minimisation  $\min_t (t - z)^2 + 2q(|t|)$  en tout  $z$  où  $\delta$  est continue. La pénalité  $q$  associée à  $\delta$  est donnée pour tout  $x \geq 0$  par*

$$q(x) = \int_0^x (r(z) - z) dz, \quad (6.7)$$

*$r$  désignant l'inverse généralisé de  $\delta$  :  $r(x) = \sup\{z \mid \delta(z) \leq x\}$ .*

Une fonction SSBS  $\delta_{\tau,\lambda}$  satisfait aux hypothèses de la Proposition 4. Il en résulte que le seuillage obtenu en utilisant une fonction  $\delta_{\tau,\lambda}$  peut être vu comme étant une approximation par régularisation avec une fonction de pénalité positive. Le résultat suivant précise la fonction de pénalité associée à  $\delta_{\tau,\lambda}$ .

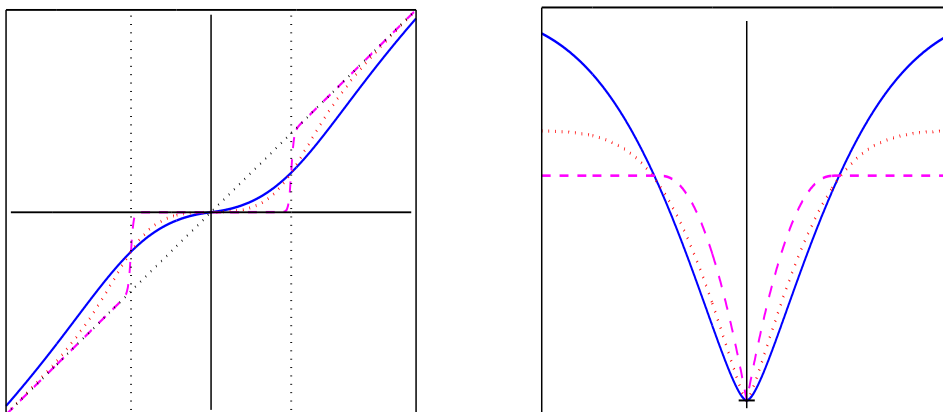
**Proposition 5** *Le seuillage obtenu avec une fonction SSBS  $\delta_{\tau,\lambda}$  est une approximation par régularisation avec la fonction de pénalité  $q_{\tau,\lambda}$ , définie pour tout  $x \geq 0$  par*

$$q_{\tau,\lambda}(x) = \frac{1}{\tau} \int_0^x \mathcal{L}(\tau z e^{-\tau(z-\lambda)}) dz, \quad (6.8)$$

*la notation  $\mathcal{L}$  désignant la fonction de Lambert définie comme étant l'inverse de la fonction :  $t \geq 0 \mapsto te^t$ .*

**Preuve :** [Atto, Pastor, and Mercier, 2008b]. ■

À partir de la Proposition 5, on obtient que pour tout  $x$  réel, la valeur  $\delta_{\tau,\lambda}(x)$  est l'unique solution du problème de minimisation  $\min_{\mathbf{d}} (x - \mathbf{d})^2 + 2q_{\tau,\lambda}(|\mathbf{d}|)$  où  $q_{\tau,\lambda}$  est donné par l'Eq. (6.8). L'allure de la fonction de pénalité  $q_{\tau,\lambda}(|x|)$  de la fonction SSBS est donnée par la figure 6.5 (à droite), pour différentes valeurs de  $\tau$ , la valeur de  $\lambda$  étant fixe. Les fonctions de pénalité représentées sur cette figure sont celles associées aux fonctions SSBS de la figure 6.5 (gauche). On



**Figure 6.5** — Allures des fonctions SSBS (à gauche), et de leurs fonctions de pénalité associées (à droite), pour différentes valeurs du degré d'atténuation  $\theta$ , avec  $\theta = \pi/6$  pour la courbe en continu (en bleu),  $\theta = \pi/4$  pour la courbe en pointillés (en rouge), et  $\theta = \pi/3$  pour la courbe en point-trait (en magenta).

constate que les pénalités associées aux SSBS sont régulières en 0, et que leur régularité en 0 dépend du degré d'atténuation. Notons la différence avec les pénalités associées aux fonctions de seuillage dur, doux, NNG et SCAD qui sont toutes singulières en 0 compte tenu du seuillage (les allures de ces fonctions sont données dans [Antoniadis, 2007, Section 3.2]). On observe également en comparant les figures 6.5 (droite) et [Antoniadis, 2007, Figure 3], que la pénalité associée à la fonction SSBS devient de plus en plus proche de celle associée à la fonction de seuillage dur lorsque le degré d'atténuation augmente. De la même manière, les degrés d'atténuation relativement faibles conduisent à des pénalités plus proches de celle de la fonction de seuillage doux.

## 6.4 Sélection des paramètres des fonctions SSBS

### 6.4.1 Optimisation SURE des paramètres des fonctions SSBS

Considérons l'estimation non-paramétrique décrite dans la section 4. La fonction de coût (risque) associée pour mesurer la précision d'un estimateur  $\hat{f}$  de  $f$  est l'EQM standard. Puisque

la transformée  $\mathcal{W}$  est orthonormale, ce coût est

$$r_\delta(\mathbf{d}, \hat{\mathbf{d}}) = \frac{1}{N} \mathbb{E} \|\mathbf{d} - \hat{\mathbf{d}}\|^2 = \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left( d_i - \delta(c_i) \right)^2 \quad (6.9)$$

pour une fonction d'atténuation  $\delta$ . L'approche SURE [Stein, 1981] est basée sur le calcul d'une estimée non-biaisée du risque  $r_\delta(\mathbf{d}, \hat{\mathbf{d}})$ . L'optimisation SURE consiste alors à chercher l'ensemble de paramètres qui minimisent cette estimée non-biaisée. Le résultat suivant est une conséquence de [Luisier, Blu, and Unser, 2007, Theorem 1].

**Proposition 6** *La quantité  $\vartheta_{\mathbf{d}} + \|\mathbf{d}\|_{\ell_2}^2 / N$ , avec  $\|\cdot\|_{\ell_2}$  désignant la norme  $\ell_2$  et*

$$\vartheta_{\mathbf{d}}(\tau, \lambda) = \frac{1}{N} \sum_{i=1}^N \frac{2\sigma^2 - d_i^2 + 2(\sigma^2 + \sigma^2 \tau |d_i| - d_i^2) e^{-\tau(|d_i| - \lambda)}}{(1 + e^{-\tau(|d_i| - \lambda)})^2}, \quad (6.10)$$

*est une estimée non-biaisée du risque  $r_{\delta_{\tau, \lambda}}(\mathbf{d}, \hat{\mathbf{d}})$ , où  $\delta_{\tau, \lambda}$  est une fonction SSBS.*

**Preuve :** [Atto, Pastor, and Mercier, 2008b]. ■

Une illustration de l'optimisation SURE pour les paramètres des fonctions SSBS est donnée ci-dessous pour l'image de 'Lena' corrompue par du bruit additif blanc et Gaussien. Dans cet exemple, nous avons considéré une TOD bi-dimensionnelle avec une ondelette 'sym8'. Les paramètres (degré d'atténuation et seuil) qui minimisent le risque (au sens SURE) associé à l'atténuation SSBS pour cet exemple sont obtenus à partir de l'algorithme du recuit simulé [Kirkpatrick, Gelatt, and Vecchi, 1983] sur l'ensemble des coefficients de détails de la TOD (paramètres globaux à appliquer sur tous les coefficients de détails). Les paramètres optimaux - au sens SURE - dépendant du niveau de décomposition (SURE LD-SSBS, LD pour Level-Dependent) sont obtenus par application de l'algorithme d'optimisation SURE aux sous-ensembles de coefficients de détails (Hori. pour Horizontal, Vert. pour Vertical, Diag. for Diagonal) obtenus à chaque niveau de décomposition (4 niveaux de décomposition). Les tables 6.1 et 6.2 donnent les valeurs des paramètres SURE SSBS et SURE LD-SSBS obtenus lorsque la minimisation du risque SURE est effectuée avec l'algorithme du recuit simulé. Ces paramètres sont donnés en terme de seuil  $\lambda$  et de degré d'atténuation  $\theta$  (compte tenu de la correspondance donnée par l'Eq. (6.5) entre  $\tau$  et  $\theta$ ). La table 6.3 résume les PSNRs obtenus lorsque l'image 'Lena' est débruitée en utilisant les paramètres optimaux SURE SSBS et SURE LD-SSBS.

**Tableau 6.1** — Paramètres SSBS optimaux au sens SURE pour la TOD basée sur une ondelette 'sym8' lorsque l'image 'Lena' est corrompue par du bruit additif blanc Gaussien. Les paramètres ont été calculés à partir de l'algorithme du recuit simulé.

$\sigma$	5	15	25	35
$\theta$	0.3064	0.5246	0.5743	0.7864
$\lambda/\sigma$	2.3833	3.1189	3.3194	3.3257

À partir des résultats de la table 6.1, on observe que le seuil, de même que le degré d'atténuation (angle  $\theta$ ), tendent à être des fonctions croissantes de l'écart-type du bruit. La même

**Tableau 6.2** — Paramètres SSBS et LD-SSBS optimaux au sens SURE et calculés par niveau de décomposition de la TOD. La TOD est basée sur une ondelette ‘sym8’ et la décomposition concerne l’image ‘Lena’ corrompue par du bruit additif blanc Gaussien. Les paramètres SSBS ont été calculés à partir de l’algorithme du recuit simulé.

$\sigma = 5$						
	$\theta$			$\lambda/\sigma$		
	Hori.	Vert.	Diag.	Hori.	Vert.	Diag.
$J = 1$	0.2866	0.2691	0.3187	3.0376	2.4020	4.1902
$J = 2$	0.2621	0.2636	0.2971	1.9457	1.7289	2.1434
$J = 3$	0.2619	0.2626	0.2621	1.5111	1.2373	1.4924
$J = 4$	0.2619	1.0433	1.0466	1.0698	1.4758	1.5565
$\sigma = 15$						
	$\theta$			$\lambda/\sigma$		
	Hori.	Vert.	Diag.	Hori.	Vert.	Diag.
$J = 1$	0.5501	0.4506	0.6868	4.7098	4.2033	4.7206
$J = 2$	0.3913	0.3761	0.4515	2.9629	2.4535	3.0973
$J = 3$	0.2671	0.2631	0.3514	2.0169	1.5666	2.1580
$J = 4$	0.2618	1.0472	1.0459	1.4356	1.0086	1.8444
$\sigma = 25$						
	$\theta$			$\lambda/\sigma$		
	Hori.	Vert.	Diag.	Hori.	Vert.	Diag.
$J = 1$	0.6833	0.5810	1.0471	4.7171	4.7180	3.7184
$J = 2$	0.4689	0.4297	0.4888	3.5136	2.7681	3.8608
$J = 3$	0.3513	0.2629	0.4067	2.4532	1.7717	2.4909
$J = 4$	0.2619	1.0469	1.0451	1.5717	1.3510	1.9566
$\sigma = 35$						
	$\theta$			$\lambda/\sigma$		
	Hori.	Vert.	Diag.	Hori.	Vert.	Diag.
$J = 1$	1.0451	0.6654	1.0471	4.0170	4.6878	4.5708
$J = 2$	0.4822	0.4292	0.5455	4.0779	3.1987	4.4296
$J = 3$	0.3990	0.2644	0.4533	2.7051	1.9608	2.7183
$J = 4$	1.0420	1.0455	0.4529	1.3756	1.2219	2.1411

remarque vaut pour l’approche dépendant du niveau de décomposition, comme on peut le vérifier à partir des résultats de la table 6.2. Cette table illustre également le fait que, pour toute valeur de  $\sigma$  considérée, le degré d’atténuation, ainsi que le seuil, tendent à devenir plus petits lorsque le niveau de décomposition augmente. En effet, l’impact du bruit devient de plus en plus faible avec le niveau de décomposition puisque seuls les coefficients d’approximation sont décomposés.

On peut aussi noter (à partir de la table 6.3) la bonne performance de la LD-SSBS (différence de moins de 1dB) en comparaison de la méthode SURELET développée dans



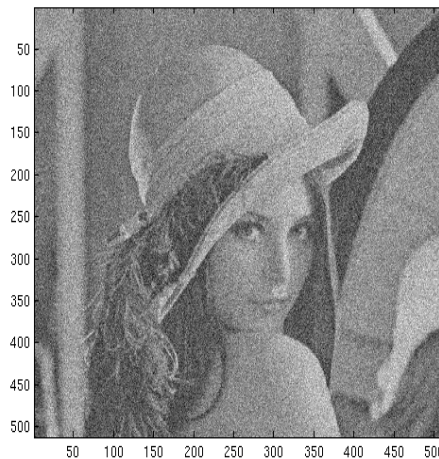
**Tableau 6.3** — PSNRs calculés lors du débruitage de l'image de 'Lena' corrompue par du bruit additif blanc Gaussien. L'approche SURE a été utilisée. Les paramètres SSBS et LD-SSBS ayant servi au calibrage de la fonction SSBS sont ceux donnés par les tables 6.1, 6.2 respectivement.

$\sigma$	5	15	25	35
input PSNR	34.15	24.61	20.17	17.25
PSNR(SSBS)	37.30	31.34	28.81	27.24
PSNR(LD-SSBS)	37.53	32.03	29.70	28.26
PSNR(SURELET)	37.96	32.68	30.36	28.88

[Luisier, Blu, and Unser, 2007]. La méthode SURELET<sup>1</sup> est également une approche SURE qui utilise une paramétrisation de l'atténuation comme somme de processus élémentaires pondérés. Cette méthode introduit en plus de l'information *a priori* sur la position des coefficients significatifs à partir d'un prédicteur inter-échelle. Les performances réalisées par les fonctions SSBS sont donc remarquables puisqu'il s'agit d'appliquer une simple fonction d'atténuation sans tenir compte d'information supplémentaire pour la localisation des coefficients significatifs.

La figure 6.7 (a-c) montre le débruitage obtenu pour l'image de 'Lena' en utilisant l'atténuation SSBS selon l'approche SURE (figure 6.7 (a-b)) et l'algorithme SURELET (figure 6.7 (c)). L'image bruitée considérée est celle de la figure 6.6.

Image bruitée, PSNR=17.25dBs



**Figure 6.6** — Image 'Lena' corrompue par du bruit additif blanc Gaussien d'écart-type  $\sigma = 35$ .

## 6.4.2 Choix heuristique des paramètres des fonctions SSBS

Cette section traite de la sélection des paramètres SSBS à travers des solutions qui ne dépendent pas d'estimée SURE de l'EQM, et donc sans qu'il soit nécessaire de passer par des

1. Code MatLab disponible via <http://bigwww.epfl.ch/demo/suredenoising/>



**Figure 6.7** — Débruitage de l'image 'Lena' corrompue par du bruit additif blanc Gaussien d'écart-type  $\sigma = 35$  par les méthodes BLS-GSM, SURELET et SSBS. Les niveaux de décomposition concernés sont  $j=1, 2, 3, 4$ .

algorithmes d'optimisation. Les techniques d'optimisation telles que le recuit simulé sont en effet assez coûteuses en calculs lorsque le nombre d'échantillons est grand. D'autre part, un risque plus faible n'est pas forcément synonyme d'augmentation de la qualité visuelle en débruitage d'images (il suffit de comparer la figure 6.7 (d) à la figure 6.7 (e) pour le constater). C'est en ce sens que, lorsqu'on manipule des images, il est souvent pratique de régler certains paramètres "à la main" pour pouvoir obtenir un niveau de qualité visuelle donné.

Cette sélection de paramètres est qualifiée d'heuristique car nous disposons de connaissances qui permettent de choisir des valeurs raisonnables pour les paramètres de la fonction SSBS. Il faut également noter que cette sélection permet de s'affranchir de la nature de la transformée en ondelettes utilisée, ce qui n'est pas le cas pour l'estimée SURE (calculée dans le cas d'une transformée orthonormale). De ce fait, et compte tenu des bonnes propriétés de la SWT en débruitage (voir section 5.3), nous allons considérer cette dernière transformée par la suite.

Notons que le degré d'atténuation  $\theta$ ,  $0 < \theta < \arccos(\sqrt{5}/5)$  donne une indication sur la régularité du processus d'atténuation. Les petites valeurs de  $\theta$  conduisent à une fonction d'atténuation très régulière. Mais dans de tels cas, les petits coefficients subissent une atténuation plus faible, d'où une moindre réduction de bruit. La sélection du degré d'atténuation peut donc être vue comme étant un compromis entre la régularité de l'atténuation et la réduction du bruit *via* le degré d'atténuation imposé aux petits coefficients.

Nous considérons pour les tests effectués dans cette section, l'image 'Lena' corrompue par du bruit additif blanc Gaussien et la SWT de cette image basée sur une ondelette de Haar avec 4 niveaux de décompositions.

Nous observons deux stratégies pour la sélection du degré d'atténuation : une valeur unique  $\theta_0 = \pi/K_0$ ,  $K_0$  étant une constante réelle telle que

$$K_0 > 2.8376 \equiv \pi / \arccos(\sqrt{5}/5);$$

ou des valeurs  $\theta(j) = \pi/(K_0 + j - 1)$  dépendant du niveau de décomposition (puisque la table 6.2 suggère l'utilisation de telles stratégies). Notons que la table 6.2 suggère aussi l'utilisation de petites (resp. grandes) valeurs du degré d'atténuation lorsque le niveau de bruit est faible (resp. fort). Par la suite, nous utiliserons les valeurs  $K_0 = 6$  lorsque l'écart-type du bruit est inférieur ou égal à 20, et  $K_0 = 4$  lorsque l'écart-type du bruit est supérieur à 20. On peut vérifier, à partir des résultats des tables 6.1 et 6.2, que les valeurs des degrés d'atténuation optimaux au sens SURE sont de cet ordre. En ce qui concerne le choix des seuils, nous utiliserons les seuils universel [Donoho and Johnstone, 1994], minimax [Donoho and Johnstone, 1994], et de détection [Atto, Pastor, and Mercier, 2008a].

### SSBS calibrée avec les seuils universel et minimax

On vérifie facilement que la fonction SSBS donne de bonnes performances lorsqu'elle est ajustée avec les seuils minimax et universel [Donoho and Johnstone, 1994] (voir figures 6.7 (d-f),  $SSBS_{\theta,\lambda}$  désignant la fonction SSBS dont les paramètres sont  $\theta$  et  $\lambda$  dans ces figures). Pour  $\theta = \pi/6$ , la suppression du bruit est beaucoup plus effective lorsque l'image est débruitée en utilisant le seuil universel, tandis que la texture et les contours sont mieux restitués au niveau de l'image débruitée avec le seuil minimax (voir figure 6.7 (d-f)).

La même remarque est vérifiée en ce qui concerne le degré d'atténuation : la texture est

mieux préservée avec l'angle  $\theta = \pi/6$ , tandis que le bruit est plus réduit lorsque  $\theta = \pi/4$  (degré d'atténuation élevé).

Les résultats expérimentaux résumés dans la table 6.4 montrent que les fonctions SSBS sont très performantes en terme de PSNR, en comparaison avec les fonctions de seuillage classiques. Cette table présente les PSNRs moyens obtenus pour les fonctions SSBS, de même que les fonctions de seuillage dur, doux, NNG et SCAD, dont les graphes sont représentés à la figure 4.4. Chaque PSNR moyen donné dans cette table a été obtenu à partir de tests basés sur 10 réalisations du bruit. Ces tests concernent l'image 'Lena' corrompue par du bruit additif blanc Gaussien d'écart-type  $\sigma = 5, 15, 25$ , et  $35$ .

**Tableau 6.4** — PSNRs moyens calculés à partir de tests sur 10 réalisations du bruit additif blanc Gaussien.

Méthode	$\sigma = 5$		$\sigma = 15$	
	$\lambda = \lambda_u$	$\lambda = \lambda_m$	$\lambda = \lambda_u$	$\lambda = \lambda_m$
Dur $_{\lambda}$	35.367	36.307	30.408	31.500
Doux $_{\lambda}$	31.570	32.969	27.030	28.198
NNG $_{\lambda}$	33.759	35.157	28.564	29.921
SCAD $_{\lambda}$	33.031	34.495	27.867	29.211
SSBS $_{\pi/6,\lambda}$	36.343	37.380	31.094	31.825
SSBS $_{\pi/4,\lambda}$	35.769	36.901	30.683	31.940

Méthode	$\sigma = 25$		$\sigma = 35$	
	$\lambda = \lambda_u$	$\lambda = \lambda_m$	$\lambda = \lambda_u$	$\lambda = \lambda_m$
Dur $_{\lambda}$	28.182	29.309	26.829	27.894
Doux $_{\lambda}$	25.334	26.337	24.383	25.265
NNG $_{\lambda}$	26.496	27.716	25.302	26.408
SCAD $_{\lambda}$	25.870	27.048	24.723	25.781
SSBS $_{\pi/6,\lambda}$	28.749	29.116	27.302	27.285
SSBS $_{\pi/4,\lambda}$	28.396	29.633	27.011	28.149

### SSBS calibrée avec les seuils de détection

Dans cette section, nous considérons les seuils de détection

$$\lambda_D(a_j, p_j^*) = \sigma \xi(a/\sigma, p^*),$$

avec  $\xi$  donné par l'Eq. (5.3) ;  $p_j^*$  et  $a_j$  sont donnés par les Eqs. (5.12) et (5.13) respectivement. Ces seuils sont plus adaptés à la nature de la décomposition en ondelettes que le seuil de détection "grossier"  $\lambda_D(N)$  (voir Section 5.3.3). L'équation (5.12) est utilisée avec la valeur  $\nu = 2.35$ , qui a été validée après quelques tests préliminaires de mise au point.

Notons que le fait d'utiliser une stratégie dépendant du niveau de décomposition va introduire des singularités car l'on utilise des fonctions SSBS différentes aux différents niveaux de décompositions concernés. Cependant, de telles stratégies conduisent à de meilleurs résultats en terme de PSNR. Plus précisément, nous allons montrer qu'une simple atténuation SSBS permet d'obtenir des performances en terme de PSNR comparables à celles d'algorithmes complexes

de débruitage d'images comme le BLS-GSM [Portilla, Strela, Wainwright, and Simoncelli, 2003]. Le BLS-GSM<sup>2</sup> est un algorithme de traitement d'images qui modélise les coefficients d'ondelettes par des vecteurs Gaussien multipliés par des scalaires positifs. C'est donc une approche paramétrique qui utilise également l'orientation et la dépendance inter-échelle des coefficients d'ondelettes. La transformée en ondelettes utilisée est une transformée redondante. La méthode a une grande complexité.

La table 6.5 présente les PSNRs obtenus, d'une part avec la fonction SSBS de paramètres  $(\theta_0, \lambda_D(a_j, p_j^*))_{j=1,2,3,4}$ , d'autre part avec l'algorithme BLS-GSM. Notons que les résultats obtenus par la fonction SSBS de paramètres  $(\theta(j), \lambda_D(a_j, p_j^*))_{j=1,2,3,4}$  sont similaires à ceux obtenus par la fonction  $SSBS_{\theta_0, \lambda_D(a_j, p_j^*)}$  et ne sont donc pas recopiés dans cette table. On constate aussi, d'après les résultats de la table 6.5, une différence de moins de 1 dB entre les PSNRs des débruitages effectués par l'algorithme BLS-GSM et la fonction SSBS ajustée avec les seuils de détection, ce qui est remarquable car la fonction SSBS telle que présentée dans ce travail n'inclut pas de prédicteurs intra / inter - échelles, ni d'autres connaissances *a priori* sur la nature de l'image analysée.

La figure 6.7 (g-h) présente les images débruitées obtenues en utilisant les seuils de détection, l'image bruitée étant celle de la figure 6.6. On a  $\sigma = 35$  pour cette image et nous avons donc fixé  $K_0 = 4$ . L'image débruitée par le BLS-GSM est donnée par la figure 6.7 (i).

En utilisant les seuils de détection, on observe un gain en PSNR et en restauration de texture, en comparaison avec les résultats réalisés par les  $SSBS_{\pi/K_0, \lambda_u}$  et  $SSBS_{\pi/K_0, \lambda_m}$ . La contrepartie est une introduction d'artefacts dus à la perte de la régularité du processus d'atténuation. On remarque sur les images de la figure 6.7 (g-h) qu'il n'y a pas de différence significative entre les débruitages obtenus avec la valeur globale  $\theta_0 = \pi/4$  et les valeurs  $\theta(j) = \pi/(4 + j - 1)$  : l'image débruitée avec la valeur  $\theta_0 = \pi/4$  semble un peu plus lisse tandis que l'image débruitée avec les valeurs  $\theta(j) = \pi/(4 + j - 1)$  a un PSNR un peu plus élevé (la différence n'est cependant pas significative).

**Tableau 6.5** — PSNRs calculés pour le débruitage de l'image 'Lena' corrompue par du bruit additif blanc Gaussien. La transformée SWT et la fonction  $SSBS_{\pi/K_0, \lambda_D(a_j, p_j^*)}$  ont été utilisées, avec  $K_0 = 6$  lorsque  $\sigma = 5, 15$  et  $K_0 = 4$  lorsque  $\sigma = 25, 35$ . La version redondante du BLS-GSM a été utilisée avec 3 orientations et des voisinages spatiaux  $3 \times 3$  dans la même sous-bande.

Méthode	$\sigma = 5$	$\sigma = 15$	$\sigma = 25$	$\sigma = 35$
PSNR initial	34.15	24.61	20.17	17.25
$SSBS_{\pi/K_0, \lambda_D(a_j, p_j^*)}$	37.37	32.51	30.31	28.91
BLS-GSM	38.22	33.52	31.28	29.79

## Impact de l'atténuation SSBS sur la distribution du signal

Dans la section 6.1, nous avons souligné l'impact du seuillage sur les distributions de signaux : les distributions obtenues après seuillage sont très singulières en ce sens que chacune d'elles est caractérisée par un pic en 0, et quelques rares coefficients éparpillés de part

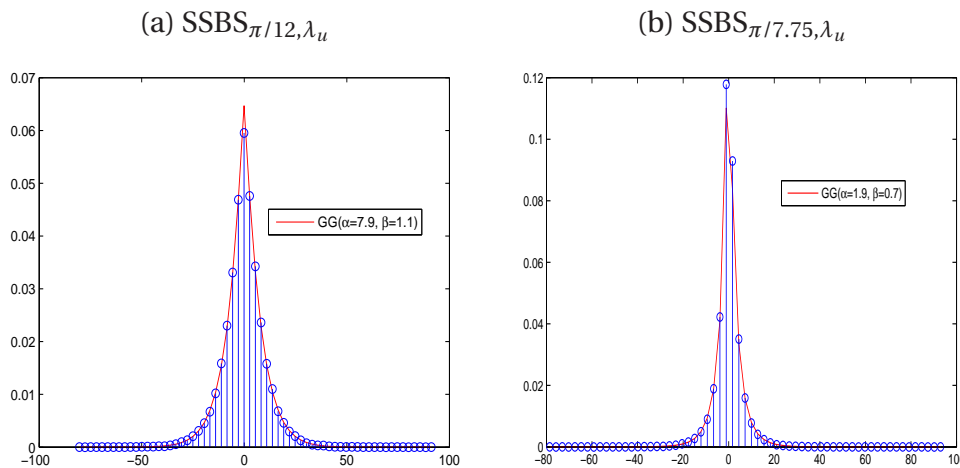
2. Code MatLab disponible via <http://decsai.ugr.es/~javier/denoise/software/index.htm/>



et d'autre de ce pic (voir figure 6.2). Ce modèle correspond à une parcimonie au sens strict et ne répond pas aux problèmes posés par les signaux réguliers par morceaux, lorsque ces derniers admettent un nombre assez élevé de singularités (ce qui est le cas des images naturelles comme celle de 'Lena', voir figure 4.2).

Nous avons vu que les fonctions SSBS permettent d'obtenir de bonnes performances, tant du point de vue PSNR que du point de vue de la qualité visuelle. Nous allons maintenant illustrer l'impact de l'atténuation SSBS sur la distribution des coefficients de l'image bruitée de 'Lena', coefficients dont l'histogramme est donné par la figure 6.1 (b). Les coefficients de l'image originale sont tels que leur histogramme est donné par la figure 6.1 (a).

L'atténuation SSBS des coefficients de l'image bruitée de 'Lena' conduisent à des distributions plus réalistes compte tenu de l'atténuation contrôlable *via* les 2 paramètres  $\theta$  et  $\lambda$  (comparer la figure 6.2 à la figure 6.8). La flexibilité des paramètres SSBS permet même d'obtenir des distributions des coefficients atténués relativement proches de la distribution de l'image originale, comme on peut le voir en comparant les figures 6.1 (a) et 6.8 (b).



**Figure 6.8** — Histogrammes obtenus après atténuation SSBS des coefficients de l'image bruitée 'Lena'. Les coefficients seuillés sont ceux représentés par l'histogramme de la figure 6.1 (b).  $\lambda_m$  est le seuil minimax. Les histogrammes sont superposés avec des distributions GG dont les paramètres sont donnés sur la figure.

## 6.5 Extension au cas du débruitage du bruit coloré

L'extension au cas du bruit additif coloré de l'estimation non-paramétrique par seuillage des coefficients des paquets d'ondelettes a été abordée dans le cadre de la sélection des seuils par certains auteurs comme [Johnstone and Silverman, 1997], [Jansen and Bultheel, 1999] ou encore [Antoniadis, Leporini, and Pesquet, 2002], lorsque les transformées en ondelettes considérées sont la TOD et ses versions redondantes comme la SWT. Notons cependant que ces transformées incluent les premières sous-bandes de décomposition, sous-bandes dont les coefficients peuvent être très corrélés lorsque le bruit est lui-même assez fortement corrélé.

D'après les Théorèmes 2 et 6, les transformées par paquets d'ondelettes,  $M$ -TPOD, décorrèlent et rendent très sensiblement Gaussiens les coefficients de processus aléatoires station-

naires (du second ordre et à bande limitée) lorsque le niveau de décomposition et la régularité des filtres sont grands. En particulier, pour l'estimation de signaux perturbés par du bruit additif coloré, et par utilisation d'une base de TPOD constituée de fonctions d'ondelettes suffisamment régulières et de bonnes résolutions, on peut se ramener très sensiblement au contexte de bruit additif blanc Gaussien ; ce qui justifie, non seulement l'utilisation de fonctions d'atténuation et des seuils standards, mais permet également de garantir des performances de l'ordre de celles obtenues dans le cas du bruit additif blanc Gaussien.

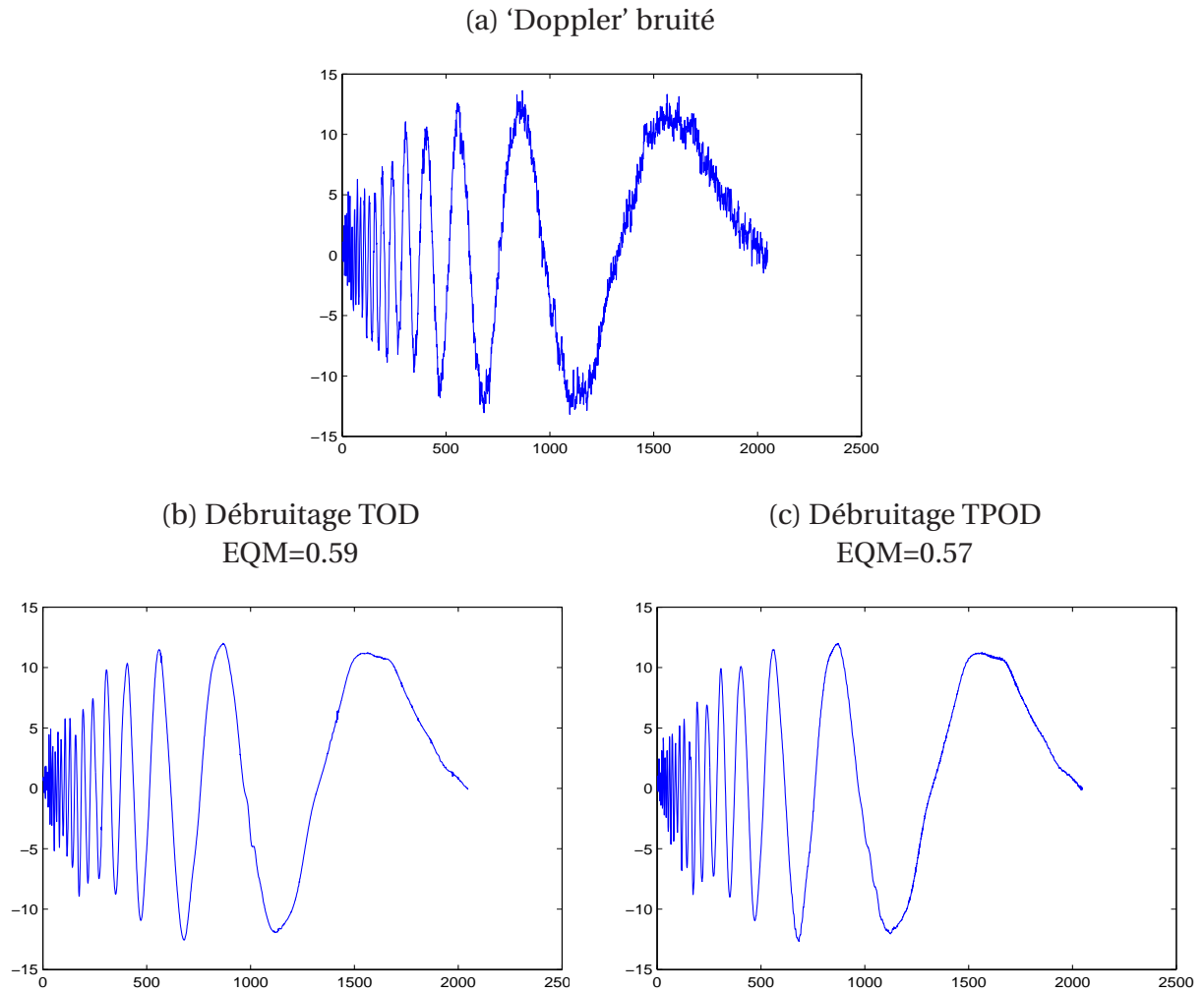
La figure 6.9 illustre l'utilisation de la TPOD pour le débruitage de signaux corrompus par du bruit additif rose (le spectre  $\gamma(\omega)$  du bruit proportionnel à  $1/\omega$ ), en comparaison avec le débruitage obtenu par la TOD. La fonction SSBS de degré d'atténuation  $\pi/5$  et de seuil universel a été utilisée pour chacune des transformées. L'ordre des filtres d'analyse est de 7. On peut ainsi noter que l'Erreur Quadratique Moyenne (EQM) obtenue avec la TPOD est plus faible que celle obtenue avec la TOD. De même en ce qui concerne la qualité visuelle du débruitage : Sur le signal débruité en utilisant la TOD, on observe un peu plus d'artefacts que sur le signal débruité en utilisant la TPOD.

**Remarque 5** *En pratique, il n'est pas nécessaire d'utiliser des filtres d'ordre élevé pour décorréler assez sensiblement les données. En effet, on obtient une très bonne capacité de décorrélation dès les premiers niveaux de décomposition lorsque le spectre du bruit n'est pas trop pointu, et cela quelque soit l'ordre choisi comme le montre l'exemple de la figure 2.3.*

*D'autre part, dans le contexte "signal + bruit coloré", le choix de l'ordre des filtres d'ondelettes doit tenir compte de la nécessité d'obtenir une représentation parcimonieuse pour le signal que l'on cherche à débruiter. En effet, il est communément admis que pour bien représenter un signal, il faut choisir une ondelette dont la régularité est aussi proche que possible de la régularité de ce signal. L'ordre du filtre devrait donc être choisi de manière à tenir compte à la fois de la capacité de décorrélation du bruit et de la capacité à ne pas trop disperser l'énergie du signal. Lorsqu'on manipule des transformées redondantes, le phénomène de moyennage induit par la redondance rend encore plus complexe la tentative de trouver l'ordre qui réalise le bon compromis. Notons que la stationnarité est perdue par utilisation des transformées non-orthogonales [Johnstone and Silverman, 1997]. Cependant, la puissance de ces transformées en débruitage est l'effet d'atténuation de bruit (effet de moyennage) induit par leur redondance.*

Un exemple d'utilisation de paquets d'ondelettes est donné à la figure 6.10 pour le débruitage de bruit coloré. Nous utilisons une version redondante (notée SWPT pour Stationary Wavelet Packet Transform) de la TPOD basée sur le principe de la SWT (pas de décimation). La décomposition utilisée est la suivante : Nous effectuons d'abord une décomposition par paquets d'ondelettes redondantes jusqu'au niveau de décomposition 3, et ensuite, nous décomposons uniquement les coefficients d'approximation obtenus aux niveaux 3, 4 et 5 selon le principe de la SWT (décomposition des chemins d'approximations à partir du troisième niveau de décomposition). Cette décomposition constitue un choix de base dans l'arbre des paquets d'ondelettes. Nous reviendrons sur le choix des bases dans les perspectives de cette thèse.

La décomposition est réalisée avec l'ondelette de Haar (Daubechies d'ordre 1), qui donne les meilleurs résultats pour les tests que nous avons effectués. Cette performance de l'ondelette de Haar n'est pas contraire aux hypothèses données ci-dessus pour le choix de l'ordre qui réalise un bon compromis entre décorrélation du bruit et capacité à fournir une représentation compacte du signal. En effet, d'une part, la redondance de la transformée introduit une complexité



**Figure 6.9** — Débruitage du signal 'Doppler' corrompu par du bruit rose. Les analyses par TOD et TPOD ont été utilisées avec une ondelette de Daubechies d'ordre 7 et 6 niveaux de décomposition. Les coefficients sont atténués par la fonction  $SSBS_{\pi/5, \lambda_u}$ ,  $\lambda_u$  étant le seuil universel. L'Erreur Quadratique Moyenne (EQM) du débruitage par TPOD est plus faible que celle obtenue avec la TOD.

supplémentaire sur le choix de la fonction d'ondelette ; d'autre part, les images naturelles sont plutôt régulières par morceaux que régulières, ce qui justifie de l'intérêt d'ondelettes à court support pour la compacité de leur représentation.

Nous avons considéré l'image 'Lena', corrompue par du bruit synthétique de type 'speckle'. Ce bruit est un bruit de type multiplicatif rencontré en général lors de l'acquisition des signaux radar, sonar ou ultrasoniques. Il peut cependant se ramener à un bruit additif mais dépendant du signal.

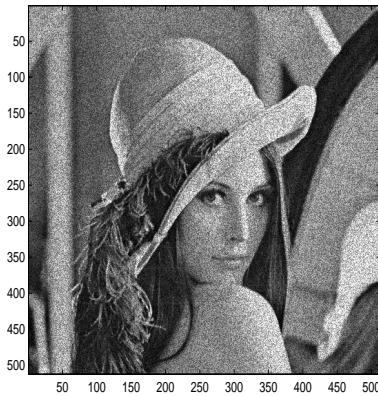
On note dans l'illustration donnée par la figure 6.10, la performance de l'estimation obtenue *via* la combinaison SWPT - SSBS (gain de l'ordre de 10 dBs comparativement au PSNR de l'image bruitée). Cette performance est de l'ordre de celles réalisées par les meilleurs algorithmes de débruitage du bruit de 'speckle' ([Xie, Pierce, and Ulaby, 2002], [Argenti and Alparone, 2002], [Achim, Tsakalides, and Bezerianos, 2003],



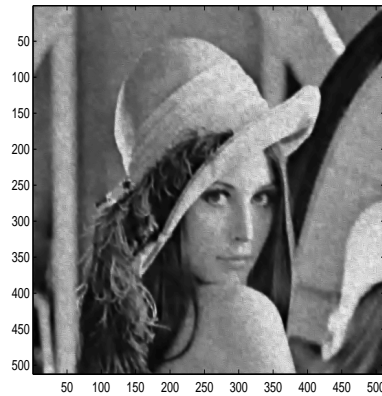
[Argenti, Bianchi, and Alparone, 2006],  
[Sen, Swamy, and Ahmad, 2007], entre autres).

[Achim, Kuruoglu, and Zerubia, 2006],

(a) Image bruitée  
PSNR = 18.83 dB



(b) SWPT - SSBS $_{\pi/5, \lambda_u}$   
PSNR = 27.84 dB



**Figure 6.10** — Débruitage de l'image 'Lena' corrompue par du bruit de type speckle. La transformée en ondelettes utilisée est une version redondante de la TPOD notée ici SWPT. La fonction SSBS est utilisée pour l'estimation des coefficients de l'image. L'atténuation est effectuée par sous-bande, en fonction de l'écart-type du bruit dans chaque sous-bande. L'écart-type du bruit est estimé par la médiane des valeurs absolues des coefficients en sous bande, normalisée par la constante 0.6745 (voir [Donoho and Johnstone, 1994]). Les filtres de décomposition sont les filtres de Haar.

---

## Conclusion

Dans cette partie, nous avons analysé l'estimation non-paramétrique de signaux par le biais d'une transformée parcimonieuse : la transformation en ondelettes. Dans un premier temps, nous avons étudié le problème du choix du seuil pour distinguer les coefficients de grandes amplitudes, ceux attribués principalement au signal, des petits coefficients ne contenant pas suffisamment d'information sur le signal et pour lesquels le bruit est prédominant. Nous avons, à partir d'une méthode non-paramétrique de détection, proposé les seuils  $\lambda_D(a, p^*)$ , dits seuils de détection, qui garantissent la même borne supérieure de la probabilité d'erreur que le test de Bayes à minimum de probabilité d'erreur pour décider si un coefficient retourné par la transformation parcimonieuse contient de l'information significative ou non. Lorsqu'ils sont utilisés pour calibrer la fonction de seuillage doux, ces seuils garantissent une borne supérieure du risque (erreur quadratique moyenne) d'estimation de deux à quatre fois plus petite que la borne supérieure du risque obtenue par utilisation des seuils minimax et universel, sur une large classe de signaux. La caractérisation de cette classe montre qu'elle est susceptible de contenir de nombreux signaux réguliers par morceaux tels que la plupart des images naturelles, cette appartenance étant fonction du niveau de bruit considéré.

Un seuil de détection dépend de deux paramètres  $a$  et  $p^*$  qui permettent de décrire le degré de parcimonie du signal étudié [Pastor and Atto, 2008]. En effet, le signal est supposé d'amplitude supérieure ou égale à une certaine valeur  $a$ , et  $p^*$  mesure une borne sur la fréquence d'apparition du signal dans l'observation. Notons que  $a$  peut être assimilée à une borne inférieure sur les amplitudes des coefficients significatifs du signal, ce qui correspond à l'amplitude minimale à partir de laquelle on veut décider que la donnée observée contient de l'information significative ou pas. Le fait que  $p^* \leq 1/2$  garantit que les coefficients significatifs du signal sont moins souvent présents qu'absent, ce qui est compatible avec l'hypothèse de parcimonie. Lorsque  $a$  est grand et  $p^*$  très petit devant  $1/2$ , on retrouve le modèle de parcimonie au sens strict décrit dans l'introduction de cette seconde partie du document. Le seuil de détection peut donc s'adapter à la parcimonie de la représentation du signal analysé par un choix très judicieux de  $a$  et  $p^*$ , lorsqu'on connaît ces valeurs ou lorsqu'il est possible de les estimer.

Ensuite, nous avons abordé le problème du choix de la fonction de seuillage. En effet, l'analyse de la parcimonie de la transformée en ondelettes sur des signaux synthétiques et des images naturelles révèle que la parcimonie découlant de la représentation en ondelettes de signaux réguliers est très fortement affectée par le nombre de singularités de tels signaux. De ce fait, le seuillage n'est pas très adéquat pour l'estimation de signaux présentant un nombre élevé de singularités, comme c'est le cas pour la plupart des images naturelles. Nous avons alors proposé de remplacer le seuillage (mise à zéro des coefficients) par une at-

ténuation plus régulière des coefficients d'ondelettes. C'est ainsi que nous avons proposé les fonctions SSBS qui permettent de contrôler, de façon flexible, l'atténuation imposée aux coefficients. Ainsi, on peut adapter le degré d'atténuation imposé aux coefficients en fonction de la nature du signal analysé et en fonction du rapport signal - à - bruit : pour un signal donné, plus la variance du bruit additif est élevée, plus forte sera l'atténuation imposée aux données. Cette adaptation du choix de paramètres en fonction du signal analysé et de la variance du bruit est d'ailleurs confirmée par une approche SURE basée sur une estimée non biaisée du risque. Le but de cette méthode est de trouver des paramètres optimaux qui minimisent le risque pour l'estimée SURE calculée à partir d'une fonction d'atténuation SSBS donnée.

Compte tenu de la régularité des fonctions SSBS, l'atténuation est effectuée avec moins de variabilité entre coefficients d'amplitudes voisines, de sorte que les propriétés statistiques du signal sont mieux conservées, et le débruitage est réalisé quasiment sans artefacts, ce qui est primordial dans certaines applications telles que l'imagerie médicale. Notons que cette propriété est également très importante lorsque l'estimation est suivie de post-traitements tels que la segmentation. Les tests effectués sur des images montrent que le débruitage obtenu avec les fonctions SSBS est un bon compromis entre la réduction de bruit et la préservation des textures et contours. Les performances réalisées par les fonctions SSBS en terme de PSNR sont de l'ordre de celles obtenues avec des algorithmes complexes de débruitage d'images tels que le BLS-GSM [Portilla, Strela, Wainwright, and Simoncelli, 2003] en comparaison duquel, le coût de calcul dû à la fonction SSBS est insignifiant.

---

# Conclusion Générale et Perspectives

## Conclusion Générale

Cette thèse apporte des contributions de deux ordres. On distingue une première contribution portant sur l'étude des propriétés statistiques des décompositions en paquets d'ondelettes par des résultats sur les fonctions d'autocorrélation et les distributions asymptotiques des coefficients de processus aléatoires stationnaires. L'autre contribution vise l'amélioration des performances en estimation non-paramétrique, par l'introduction des seuils de détection et des fonctions d'atténuation SSBS, afin de mieux prendre en compte la nature parcimonieuse des représentations en ondelettes des signaux réguliers et réguliers par morceaux.

## Propriétés statistiques des transformations en ondelettes

Les résultats énoncés dans la première partie du document concernent un processus aléatoire  $X$ , supposé être stationnaire, du second ordre, et à bande limitée. À partir du Théorème 6, on sait qu'une transformée en paquets d'ondelettes décorrèle très sensiblement et rend quasiment Gaussiens, les coefficients d'ondelettes de  $X$  lorsque le niveau de résolution  $j$  est supérieur ou égal à un certain niveau  $j_0$  donné, et pourvu que la régularité  $r$  des fonctions d'ondelettes dépasse un certain ordre  $r_0$  qui, lui-même, dépend de  $j_0$ . Pour un processus donné, et lorsqu'on parcourt les nœuds d'un chemin de l'arbre des paquets d'ondelettes, la vitesse de convergence vers le processus limite dépend à la fois de la régularité des filtres d'analyse et de l'allure des polyspectres du processus étudié.

Ainsi, la décorrélation est plus lente dans les chemins associés aux sous-bandes pour lesquelles, la pente de la densité spectrale de puissance du processus décomposé est très forte. Dans le cas d'une densité spectrale de puissance à pente relativement faible, l'autocorrélation décroît très vite avec le niveau de décomposition. On a le même constat en ce qui concerne la distribution des coefficients d'ondelettes aux nœuds du chemin parcouru : les coefficients d'un processus  $X$ , de distribution assez régulière, tendent très vite vers la distribution limite associée au chemin choisi, lorsque le niveau de résolution et la régularité des filtres de décomposition augmentent.

À l'opposé, lorsque le processus en entrée d'une décomposition en paquets d'ondelettes admet une densité spectrale avec une pente très accentuée au voisinage d'une certaine fréquence,

il est relativement difficile, dans les chemins associés aux sous bandes incluses dans ce voisinage, d'obtenir une bonne décorrélation avec des filtres de faible régularité. De même pour les processus admettant des distributions très singulières : il est nécessaire d'augmenter considérablement à la fois le niveau de décomposition et la régularité des filtres d'analyse pour obtenir une bonne appréciation de la convergence vers la distribution limite, dans les sous-bandes associées à certains chemins de l'arbre des paquets d'ondelettes. L'exploitation pratique de ces résultats est donc délicate pour des processus admettant de très fortes corrélations ou de très singulières distributions. En effet, l'augmentation de la régularité des filtres d'ondelettes se traduit par de l'instabilité numérique compte tenu du fait que les filtres à forte régularité sont des filtres numériques à longues réponses impulsionnelles.

Notons enfin que ces résultats sur les propriétés statistiques des décompositions en ondelettes et par paquets d'ondelettes précisent les conditions pour lesquelles l'estimation non-paramétrique en présence du bruit additif coloré se ramène à l'estimation en présence de bruit additif quasi-décorrélé et très sensiblement Gaussien par sous bandes. Cependant, dans le contexte "signal + bruit additif coloré", il faut aussi tenir compte de la régularité qui permette de réaliser une représentation compacte du signal : D'où le compromis à faire entre parcimonie de la représentation du signal et capacité de décorrélation du bruit pour le choix de l'ordre des filtres d'analyse. Pour ces raisons, il semble préférable d'effectuer une recherche de la meilleure base de paquets d'ondelettes qui réalise une décorrélation acceptable sans trop disperser l'énergie du signal, donc avec des valeurs pas très élevées pour les niveaux de résolution et la régularité des fonctions d'ondelettes. Cette analyse est reportée dans les perspectives qui sont données par la suite.

## Estimation non-paramétrique de signaux

La seconde contribution de ce travail de recherche a porté sur l'amélioration des performances en estimation non-paramétrique de signaux, par l'introduction des seuils de détection et des fonctions d'atténuation SSBS. L'objectif de ces approches est de proposer des traitements qui tiennent compte du degré de parcimonie des représentations en ondelettes des signaux réguliers et réguliers par morceaux. Nous avons ainsi synthétisé une famille de fonctions d'atténuation, dites fonctions SSBS (Smooth Sigmoid Based Shrinkage), qui sont régulières, à degré d'atténuation flexible, et qui permettent de réduire le bruit blanc Gaussien sans introduire d'artefact de reconstruction. Les fonctions SSBS, utilisés avec une transformée en ondelettes invariante par translation, réalisent de très bonnes performances en comparaison avec les meilleurs algorithmes de débruitage d'images. Ces fonctions, lorsqu'elles sont ajustées avec les seuils universels, minimax ou de détection, peuvent aussi être utilisés pour l'estimation de signaux recueillis en présence de bruit coloré, tout en garantissant des performances proches de celles réalisées dans le cas du bruit blanc et Gaussien. Elles jouissent, comparativement aux fonctions de seuillage classiques, d'un atout supplémentaire pour le débruitage en présence de bruit coloré : le modèle du bruit blanc et Gaussien n'étant pas formellement valide pour une représentation en ondelettes de bruit coloré (filtres de régularité finie, niveau de décomposition fini), les seuils idéaux permettant de distinguer le signal du bruit sont difficiles à évaluer, et l'usage de fonctions SSBS suffisamment régulières permet de réduire le bruit coloré avec une très bonne performance en terme de PSNR. Cela est illustré à la fin de la seconde partie du document (Section 6.5), pour le débruitage d'une image corrompue par du bruit de

speckle, et lorsque la transformée en ondelettes utilisée est invariante par translation. Une justification de cette performance vient du fait que les fonctions SSBS n'imposent pas de seuillage, mais atténuent les données de manière régulière avec un facteur d'atténuation plus fort pour les données de faibles amplitudes que pour celles qui sont de grandes amplitudes.

## Perspectives

### Propriétés statistiques des transformations en ondelettes

Nous avons vu que la  $M$ -TPOD réalise une discrétisation quasi-optimale car elle permet de transformer des processus stationnaires en séquences de données quasi-décorrélées et sensiblement Gaussiennes. De surcroît, le fait que *le spectre du processus d'entrée distribue ses valeurs aux variances des paquets d'ondelettes associés aux divers chemins de la  $M$ -TPOD* permet d'envisager de nombreuses applications. En particulier, nous pensons que ces résultats peuvent contribuer à l'étude de l'échantillonnage non-uniforme de la densité spectrale d'un signal (voir [Yu and Karlsson, 2001], [Cristán and Walden, 2002], entre autres). En effet, la suite des variances associées aux sous-bandes d'un arbre de paquets d'ondelettes constitue une forme de *discrétisation de spectre* lorsque la décorrélation est atteinte. L'étude de la variance des paquets d'ondelettes en vue de la détection de changement semble être une autre application d'intérêt ([Gabbanini, Vannucci, Bartoli, and Moro, 2004], entre autres).

Une extension immédiate concernant les propriétés statistiques des transformations par paquets d'ondelettes discrètes est la recherche de la meilleure base au sens du critère de blancheur locale ou au sens de la normalité des séquences de coefficients, compte tenu du fait que la décorrélation et la tendance à la normalité ne sont pas homogènes aux différents nœuds  $(j, n)$ , pour  $j$  fixé, et lorsque  $n = 0, 1, 2, \dots, M^j - 1$ . L'idée est de choisir une régularité raisonnable pour les filtres d'analyse et de ne pas décomposer davantage les coefficients ayant atteint un certain seuil de décorrélation préalablement fixé. Cette extension a deux avantages : d'une part, elle réduira la complexité algorithmique de la transformée ; d'autre part, elle permettra d'obtenir une représentation mieux localisée dans le domaine fréquentiel (la mauvaise localisation fréquentielle de la décomposition complète par paquets d'ondelettes a pour effet de réduire les performances de certains traitements tels que le débruitage). Cette analyse pourra se faire en comparaison ou de manière complémentaire à l'algorithme de recherche de la meilleure base fondé sur un critère entropique [Coifman and Wickerhauser, 1992]. Son intérêt est qu'elle peut permettre de justifier certains compromis couramment utilisés en compression<sup>3</sup> et en modulation<sup>4</sup> par ondelettes.

Une autre perspective sur l'étude des propriétés statistiques des transformations en ondelettes et par paquets d'ondelettes concerne le cas de certains bruits non-stationnaires, tel que ceux induits par les processus cyclostationnaires, et l'impact de la régularité des filtres d'ondelettes sur les statistiques des coefficients de ces bruits.

3. adapter la régularité des filtres d'ondelettes à la nature de l'image à compresser conduit à un bon compromis entre le taux de compressibilité important et une qualité acceptable pour l'image compressée [Kharate, Patil, and Bhale, 2007].

4. les filtres à fortes régularité réduisent l'interférence entre porteuses, mais augmentent l'interférence entre symboles, dans la modulation en ondelettes [Negash and Nikookar, 2001].



D'autre part, la convergence des filtres d'ondelettes en  $M$ -bandes de Daubechies et de Battle-Lemarié vers les filtres de Shannon, pour  $M \geq 3$ , n'a pas encore été formellement établie (après une recherche bibliographique qui est difficilement exhaustive, de nos jours). Une telle convergence, si elle est prouvée, est de nature à accélérer les convergences évoquées par les Théorèmes 2 et 6 portant respectivement sur les fonctions d'autocorrélation et les distributions des coefficients des paquets d'ondelettes, puisque les petites valeurs de  $M$  sont les moins favorables.

Enfin, notons que la synthèse de familles de filtres d'ondelettes dont les transformées de Fourier convergent presque partout vers les transformées de Fourier des filtres de Shannon est une perspective d'intérêt au vu des résultats donnés par les Théorèmes 2 et 6.

## Seuils de détection

L'une des perspectives les plus importantes au sujet de la caractérisation des seuils de détection est l'analyse de la régularité qu'ils induisent selon l'appartenance de l'estimée à un espace de Besov donné, en connexion avec les résultats énoncés dans [Donoho and Johnstone, 1996, 1999]. En particulier, cette analyse peut conduire à des sous-classes de seuils de détection en fonction des bornes sur les erreurs quadratiques moyennes qu'ils engendrent, et selon la régularité Besovienne de la fonction à estimer. Notons à ce niveau que les seuils minimax,  $\lambda_m(N)$ , et universels,  $\lambda_u(N)$ , sont des seuils de détection qui correspondent à des valeurs spécifiques des bornes  $p^*$  et  $a$ , qui sont des bornes - supérieures sur la probabilité de présence - inférieure sur l'amplitude minimale significative - du signal, lorsque ce dernier est observé en présence de bruit blanc Gaussien. La résolution de l'Eq.  $\lambda_D(a, p^*) = \lambda_u(N)$  en fonction de  $a$  et  $p^*$  définit ainsi une surface dont les points sont des indices de parcimonie. De même pour l'Eq.  $\lambda_D(a, p^*) = \lambda_m(N)$ . La caractérisation de ces surfaces ou des sections de ces surfaces ( $a$  fixé ou alors  $p^*$  fixé) n'a pas encore été abordée dans ce travail, mais va dans le même sens que l'analyse et la classification des seuils de détection en fonction de la régularité du signal analysé.

Les seuils de détection mesurent la parcimonie d'un signal à travers les paramètres  $p^*$  et  $a$ . Ces paramètres sont étroitement liés à la distribution du signal étudié. L'extraction de ces paramètres à partir de la connaissance *a priori* de la distribution du signal peut permettre d'améliorer les performances des seuils de détection pour l'estimation de signaux.

## Estimation par SSBS

Les tests de débruitage basés sur les fonctions SSBS ont été effectués sur des images naturelles. Cependant, on peut espérer obtenir des résultats intéressants en débruitage de la parole compte tenu du fait que les fonctions SSBS permettent de réduire l'effet du bruit sans trop dégrader les propriétés statistiques du signal.

Les fonctions SSBS sont très flexibles et un choix judicieux des paramètres de cette famille de fonctions permet de ne pas trop affecter la distribution du signal. Un problème concernant le choix de ses paramètres est : Connaissant *a priori* la distribution du signal, quels paramètres choisir pour que la distribution de l'estimée par atténuation SSBS se rapproche au mieux de la distribution du signal ?

Du point de vue de la qualité de l'estimation, on peut raisonnablement espérer améliorer les

performances des SSBS par l'introduction de prédicteur inter et/ou intra-échelle concernant la position de coefficients significatifs du signal. On sait en effet que pour un signal régulier par morceaux, chaque singularité crée un certain nombre de petits coefficients à chaque niveau décomposition. Le rôle du prédicteur sera alors de détecter les singularités du signal et de localiser leurs contributions au sein des coefficients d'ondelettes. Quelques exemples de prédicteurs sont donnés dans [Portilla, Strela, Wainwright, and Simoncelli, 2003] pour le BLS-GSM et dans [Luisier, Blu, and Unser, 2007] pour l'algorithme SURELET.

Une autre extension concerne l'atténuation SSBS par blocs de coefficients [Cai, 1996]. L'idée est de calculer les paramètres SSBS (degré d'atténuation et seuil) par blocs de coefficients, en fonctions de l'énergie et de la localisation du bloc considéré.

Le dernier point visant à l'amélioration des performances SSBS concerne l'utilisation de ces fonctions avec les trois paramètres  $t$ ,  $\tau$  et  $\lambda$ , c'est à dire les fonctions  $\delta_{t,\tau,\lambda}$ , ce qui donne encore plus de flexibilité à l'atténuation des données. Cette extension vise notamment le calcul d'une estimée SURE dépendant des fonctions SSBS  $\delta_{t,\tau,\lambda}$  et une extraction des trois paramètres  $(t, \tau, \lambda)$  qui minimisent l'estimée du risque SURE calculée.





---

## Bibliographie

- A. Achim, P. Tsakalides, and A. Bezerianos. Sar image denoising via bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 41(8) :1773 – 1784, Aug. 2003.
- A. Achim, E. E. Kuruoglu, and J. Zerubia. Sar image filtering based on the heavy-tailed rayleigh model. *IEEE Transactions on Image Processing*, 15(9) :2686 – 2693, Sept. 2006.
- A. Aldroubi, M. Unser, and M. Eden. Cardinal spline filters : Stability and convergence to the ideal sinc interpolator. *Signal Process*, 28(8) :127–138, Aug. 1992.
- A. Antoniadis. Wavelet methods in statistics : Some recent developments and their applications. *Statistics Surveys*, 1 :16–55, 2007.
- A. Antoniadis and J. Fan. Regularization of wavelet approximations. *Journal of the American Statistical Association*, 96(455) :939–955, Sep. 2001.
- A. Antoniadis and G. Oppenheim. *Wavelets and Statistics*. Springer-Verlag, New York, 1995.
- A. Antoniadis, D. Leporini, and J.-C. Pesquet. Wavelet thresholding for some classes of non-gaussian noise. *Statistica Neerlandica*, 56(4) :434–453, 2002.
- F. Argenti and L. Alparone. Speckle removal from sar images in the undecimated wavelet domain. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11) :2363 – 2374, Nov. 2002.
- F. Argenti, T. Bianchi, and L. Alparone. Multiresolution map despeckling of sar images based on locally adaptive generalized gaussian pdf modeling. *IEEE Transactions on Image Processing*, 15(11) :3385 – 3399, Nov. 2006.
- A. M. Atto and D. Pastor. Limit distributions for wavelet packet coefficients of band-limited stationary random processes. *European Signal Processing Conference, EUSIPCO*, Lausanne, Switzerland, August 25-28, 2008a.
- A. M. Atto and D. Pastor. Central limit theorems for wavelet packet decompositions of stationary random processes. Submitted to *IEEE Transactions on Information Theory*, <http://aps.arxiv.org/abs/0802.0797v2>, 2008b.
- A. M. Atto, D. Pastor, and A. Isar. On the asymptotic decorrelation of the wavelet packet coefficients of a wide-sense stationary random process. *International Symposium on Electronics and Telecommunications, ETC*, Timisoara, Romania, September 21-23, 2006.

- A. M. Atto, D. Pastor, and A. Isar. On the statistical decorrelation of the wavelet packet coefficients of a band-limited wide-sense stationary random process. *Signal Processing*, 87(10) : 2320–2335, Oct. 2007.
- A. M. Atto, D. Pastor, and G. Mercier. Detection threshold for non-parametric estimation. *Signal, Image and Video Processing, Springer*, 2(3), Sept. 2008a.
- A. M. Atto, D. Pastor, and G. Mercier. Smooth adaptation by sigmoid shrinkage. Submitted to *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008b.
- A. M. Atto, D. Pastor, and G. Mercier. Smooth sigmoid wavelet shrinkage for non-parametric estimation. *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Las Vegas, Nevada, USA, 30 march - 4 april, 2008c.
- S. M. Berman. *Sojourns and extremes of stochastic processes*. Wadsworth and Brooks/Cole, 1992.
- L. Breiman. Better subset regression using the nonnegative garrote. *Technometrics*, 37 :373–384, 1995.
- A. G. Bruce and H. Y. Gao. Understanding waveshrink : Variance and bias estimation. *Biometrika*, 83(4) :727–745, 1996.
- T. T. Cai. Minimax wavelet estimation via block thresholding. Technical Report 96-41, 1996. URL [citeseer.ist.psu.edu/cai96minimax.html](http://citeseer.ist.psu.edu/cai96minimax.html).
- A. Cohen. *Ondelettes et traitement numérique du signal*. Masson, Paris, 1992.
- R. R. Coifman and D. L. Donoho. *Translation invariant de-noising*, pages 125–150. Number 103. Lecture Notes in Statistics, 1995.
- R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2) :713–718, Mar. 1992.
- P. F. Craigmle and D. B. Percival. Asymptotic decorrelation of between-scale wavelet coefficients. *IEEE Transactions on Information Theory*, 51(3) :1039–1048, Mar. 2005.
- A. C. Cristán and A. T. Walden. Multitaper power spectrum estimation and thresholding : Wavelet packets versus wavelets. *IEEE Transactions on Signal Processing*, 50(12) :2976–2986, Dec. 2002.
- L. Şendur and I. V. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on Signal Processing*, 11 :2744–2756, Nov. 2002.
- I. Daubechies. *Ten lectures on wavelets*. SIAM, Philadelphie, PA, 1992.
- R. W. Dijkerman and R. R. Mazumdar. On the correlation structure of the wavelet coefficients of fractional brownian motion. *IEEE Transactions on Information Theory*, 40(5) :1609–1612, Sep. 1994.

- M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, 11(2) :146–158, Feb. 2002.
- D. L. Donoho. Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect noisy data. *Proceedings of Symposia in Applied Mathematics*, 1993a.
- D. L. Donoho. Unconditional bases are optimal bases for data compression and for statistical estimation. *Applied and Computational Harmonic Analysis*, 1(1) :100–115, 1993b.
- D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3) :613–627, May 1995.
- D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3) :425–455, Aug. 1994.
- D. L. Donoho and I. M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432) :1200–1224, Dec. 1995.
- D. L. Donoho and I. M. Johnstone. Neo-classical minimax problems, thresholding and adaptive function estimation. *Bernoulli* 2, (1) :39–62, 1996.
- D. L. Donoho and I. M. Johnstone. Asymptotic minimaxity of wavelet estimators with sampled data. *Statistica Sinica*, 9(1) :1–32, 1999.
- P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2) :910–917, Mar. 1992.
- F. Gabbanini, M. Vannucci, G. Bartoli, and A. Moro. Wavelet packet methods for the analysis of variance of time series with application to crack widths on the brunelleschi dome. *Journal of Computational and Graphical Statistics*, 13(3) :639–658, 2004.
- H. Y. Gao. Waveshrink shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7(4), 1998.
- W. Härdle, G. Kerkycharian, D. Picard, and A. Tsybakov. *Wavelets, Approximation, and statistical Applications*. Springer-Verlag, Lecture notes in statistics, New York, 1998.
- C. Houdré. Wavelets, probability, and statistics : Some bridges. In J. J. Benedetto and M. W. Frasier, editors, *Wavelets : Mathematics and applications*, chapter 9, pages 365–398. CRC Press, 1994.
- A. Isar and I. Naforniță. *Représentations temps-frequence*. Ed Politehnica, Timișoara, Roumanie, 1998.
- M. Jansen and A. Bultheel. Multiple wavelet threshold estimation by generalized cross validation for images with correlated noise. *IEEE Transactions on Image Processing*, 8(7) :947–953, Jul. 1999.
- I. M. Johnstone. Wavelets and the theory of non-parametric function estimation. *Journal of the Royal Statistical Society*, A(357) :2475–2493, 1999.

- I. M. Johnstone and B. W. Silverman. Wavelet threshold estimators for data with correlated noise. *Journal of the Royal Statistical Society, Series B*, 59(2) :319+, 1997.
- G. K. Kharate, V. H. Patil, and N. L. Bhale. Selection of mother wavelet for image compression on basis of nature image. *Journal of the Multimedia*, 2(6) :44–51, 2007.
- S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220 (4598) :671–680, 1983.
- H. Krim, D. Tucker, S. Mallat, and D. L. Donoho. On denoising and best signal representation. *IEEE Transactions on Information Theory*, 45(7) :2225+, Nov. 1999.
- C. Lemoine and D. Pastor. Démodulation cohérente par analyse multi-résolution en présence de bruit parasite coloré et gaussien. *Traitement du Signal*, 15(3) :251–264, 1998.
- D. Leporini and J.-C. Pesquet. High-order wavelet packets and cumulant field analysis. *IEEE Transactions on Information Theory*, 45(3) :863–877, Apr. 1999.
- F. Luisier, T. Blu, and M. Unser. A new sure approach to image denoising : Interscale orthonormal wavelet thresholding. *IEEE Transactions on Image Processing*, 16(3) :593–606, Mar. 2007.
- S. Mallat. *A wavelet tour of signal processing, second edition*. Academic Press, 1999.
- M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi. *Les ondelettes et leurs Applications*. Hermes Science, Lavoisier, 2003.
- B. G. Negash and H. Nikookar. Wavelet based ofdm for wireless channels. *Vehicular Technology Conference*, Rhodes, Greece, June 06-09, 1 :688–691, 2001.
- D. Pastor and A. M. Atto. Sparsity from binary hypothesis testing and application to non-parametric estimation. *European Signal Processing Conference, EUSIPCO*, Lausanne, Switzerland, August 25-29, 2008.
- D. Pastor and R. Gay. Décomposition d'un processus stationnaire du second ordre : Propriétés statistiques d'ordre 2 des coefficients d'ondelettes et localisation fréquentielle des paquets d'ondelettes. *Traitement du Signal*, 12(5), 1995.
- D. Pastor, R. Gay, and A. Gronenboom. A sharp upper bound for the probability of error of likelihood ratio test for detecting signals in white gaussian noise. *IEEE Transactions on Information Theory*, 48(1) :228–238, Jan. 2002.
- J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image processing*, 12(11) :1338–1351, November 2003.
- N. Saito and G. Beylkin. Multiresolution representation using the autocorrelation functions of compactly supported wavelets. *IEEE Transactions on Signal Processing*, 41 :3584–3590, Dec. 1993.

- D. Sen, M. N. S. Swamy, and M. O. Ahmad. Computationally fast techniques to reduce awgn and speckle in videos. *IET Image Processing*, 1(4) :319 – 334, 2007.
- R. J. Serfling. *Approximation theorems of mathematical statistics*. John Wiley and Sons, 1980.
- J. Shen and G. Strang. Asymptotic analysis of daubechies polynomials. *Proceedings of the American Mathematical Society*, 124(12) :3819–3833, December 1996.
- J. Shen and G. Strang. Asymptotics of daubechies filters, scaling functions, and wavelets. *Applied and Computational Harmonic Analysis*, 5(HA970234) :312+, 1998.
- P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus. Theory of regular  $m$ -band wavelet bases. *IEEE Transactions on Signal Processing*, 41(12) :3497–3511, Dec. 1993.
- C. Stein. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, 9 :1135–1151, 1981.
- A. H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2) :904–909, Mar. 1992.
- S. Touati and J.-C. Pesquet. Some results on the wavelet packet decomposition of nonstationary processes. *EURASIP Journal on Applied Signal Processing*, 2002(11) :1289–1295, Nov. 2002.
- G. G. Walter. *Wavelets and other orthogonal systems with applications*. CRC Press, 1994.
- H. Xie, L. E. Pierce, and F. T. Ulaby. Sar speckle reduction using wavelet denoising and markov random field modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10) :2196 – 2212, Oct. 2002.
- J. Yu and S. Karlsson. Local spectral analysis using wavelet packets. *Circuits Systems Signal Processing*, 20(5) :497–528, 2001.
- J. Zhang and G. Walter. A wavelet-based KL-like expansion for wide-sense stationary random processes. *IEEE Transactions on Signal Processing*, 42(7) :1737–1745, July 1994.



## A

# Représentations dans les espaces de Hilbert

## A.1 Rappels d'analyse

**Définition 1 (Suite de Cauchy)** Une suite  $(x_\ell)_\ell$  d'un espace vectoriel normé, de norme  $\|\cdot\|$ , est dite suite de Cauchy si :

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N}: \quad \ell, k > N \Rightarrow \|x_\ell - x_k\| < \epsilon. \quad (\text{A.1})$$

**Définition 2 (Espace de Banach)** Un espace de Banach est un espace vectoriel normé et complet.

La *complétude* assure que toute suite de Cauchy dans un espace vectoriel normé converge vers un élément de cet espace vectoriel.

**Définition 3 (Espace de Hilbert)** C'est un espace de Banach, muni d'un produit scalaire.

Le produit scalaire, qui sera noté  $\langle \cdot, \cdot \rangle$ , donne un sens à la notion de projection (angles et orthogonalité découlent de la définition d'un produit scalaire).

**Définition 4 (Espace de Hilbert  $L^2(\mathbb{R})$ )** Une fonction  $f$  de variable réelle, et à valeur dans  $\mathbb{R}$  ou  $\mathbb{C}$ , appartient à  $L^2(\mathbb{R})$  si

$$\int_{\mathbb{R}} |f(t)|^2 dt < \infty.$$

Le produit scalaire défini sur cet espace est défini par :

$$f, g \in L^2(\mathbb{R}) \Rightarrow \langle f, g \rangle = \int_{\mathbb{R}} f(t) \overline{g(t)} dt.$$

La norme de  $f \in L^2(\mathbb{R})$  est définie par :

$$\|f\|_2 = \left( \int_{\mathbb{R}} |f(t)|^2 dt \right)^{\frac{1}{2}}. \quad (\text{A.2})$$

On notera aussi cette norme  $\|\cdot\|$  lorsqu'il n'y a pas de confusion possible.



**Définition 5 (Transformée de Fourier dans  $L^1(\mathbb{R})$ )** Soit  $f$  une fonction de  $L^1(\mathbb{R})$  (espace des fonctions de la variable réelle et absolument intégrables). La transformée de Fourier de  $f$ , qu'on notera  $\mathcal{F}f$  ou  $\hat{f}$  est définie par

$$\hat{f} = \mathcal{F}f = \int_{\mathbb{R}} f(t) e^{-i\xi t} dt. \quad (\text{A.3})$$

Si  $\hat{f} \in L^1(\mathbb{R})$ , la transformée de Fourier inverse  $\overline{\mathcal{F}\hat{f}}$  de  $\hat{f}$ , est définie par

$$f = \overline{\mathcal{F}\hat{f}} = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) e^{+i\xi t} d\xi. \quad (\text{A.4})$$

**Définition 6 (Transformée de Fourier dans  $L^2(\mathbb{R})$ )**  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  étant dense dans  $L^2(\mathbb{R})$ , la transformée de Fourier de  $f \in L^2(\mathbb{R})$  se définit comme limite de transformées de Fourier d'une suite de  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  convergeant vers  $f$ . La convergence a lieu au sens de la norme de  $L^2(\mathbb{R})$ .

**Théorème 7 (Formule de Parseval - Plancherel)** Soient  $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ , alors (formule de Parseval)

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) \overline{g(t)} dt = \frac{1}{2\pi} \langle \hat{f}, \hat{g} \rangle. \quad (\text{A.5})$$

En particulier, si  $f = g$ , on en déduit que (formule de Plancherel)

$$\|f\|_2^2 = \frac{1}{2\pi} \|\hat{f}\|_2^2. \quad (\text{A.6})$$

La régularité d'une fonction est liée, et se mesure souvent par le nombre de fois où cette fonction est différentiable. On notera  $\mathcal{C}^\ell$ , l'espace des fonctions  $\ell$ -fois continûment différentiables. Lorsqu'on s'intéresse à la régularité d'une fonction de  $L^2(\mathbb{R})$ , il est plus commode de travailler dans les espaces de Sobolev. Soit  $f \in L^2(\mathbb{R})$  une fonction suffisamment régulière et dérivable jusqu'à l'ordre  $\ell \in \mathbb{N}$ . On a :

$$\mathcal{F}\left(\frac{d^\ell}{dx^\ell} f\right) = (i\xi)^\ell \mathcal{F}f(\xi).$$

Or, d'après la formule de Parseval-Plancherel, on a :

$$\|\mathcal{F}\left(\frac{d^\ell}{dx^\ell} f\right)\|_2 = \left\| \frac{d^\ell}{dx^\ell} f \right\|_2.$$

La dérivée d'ordre  $\ell$  de  $f$  appartiendra à  $L^2(\mathbb{R})$  si  $\|\frac{d^\ell}{dx^\ell} f\|_2 < \infty$ , c'est à dire si :

$$\|\mathcal{F}\left(\frac{d^\ell}{dx^\ell} f\right)\|_2 = \int_{\mathbb{R}} |\xi|^{2\ell} |\hat{f}(\xi)|^2 d\xi < \infty.$$

**Définition 7 (Différentiabilité au sens de Sobolev)** Soit  $f \in L^2(\mathbb{R})$  et soit  $s$  un réel strictement positif. On dira que  $f$  est  $s$  fois différentiable au sens de Sobolev si

$$\int_{\mathbb{R}} |\xi|^{2s} |\hat{f}(\xi)|^2 d\xi < \infty \quad (\text{A.7})$$

La formule donnée par l'Eq. (A.7) nous dit que  $|\hat{f}(\xi)|$  décroît plus rapidement que  $|\xi|^{-2s}$ , pour  $s > 0$ . Cette définition donne un sens plus général à la notion de différentiabilité en l'étendant aux ordres  $s$  non entiers par passage dans le domaine de Fourier. Entre les fonctions  $\ell$ -fois et  $(\ell + 1)$ -fois différentiables, il existe toute une classe de fonctions (pour  $\ell < s < \ell + 1$ ) dont la régularité  $s$  est précisée par l'Eq. (A.7). En particulier,  $f$  sera différentiable au sens de Sobolev si  $f' \in L^2(\mathbb{R})$ , c'est à dire si

$$\int_{\mathbb{R}} |\xi|^2 |\hat{f}(\xi)|^2 d\xi < \infty. \quad (\text{A.8})$$

On définit alors les espaces de Sobolev par :

**Définition 8 (Espace de Sobolev  $W_2^s(\mathbb{R})$ )** C'est l'ensemble des fonctions  $f$  appartenant à  $L^2(\mathbb{R})$  et  $s$  fois différentiable au sens de Sobolev.

**Remarque 6** 1. Pour  $s = 0$ , on a :  $W_2^0(\mathbb{R}) = L^2(\mathbb{R})$ .

2. Pour  $s > n + \frac{1}{2}$ , on a :  $f \in \mathcal{C}^n$ .

3. La norme sur l'espace de Sobolev  $W_2^s$ , induite par la norme de  $L^2(\mathbb{R})$  est telle que pour  $f \in W_2^s(\mathbb{R})$ ,

$$\|f\|_{W_2^s} = \|f\|_2 + \|f^{(s)}\|_2. \quad (\text{A.9})$$

Pour aller encore plus loin dans la caractérisation de la régularité d'une fonction  $f \in L^2(\mathbb{R})$ , nous allons maintenant définir les *espaces de Besov*. Pour cela, définissons les opérateurs et les modules de continuité,

**Définition 9 (Opérateurs de continuité)** Ce sont les opérateurs  $\Delta_h^1$  et  $\Delta_h^2$  définis par :

$$\Delta_h^1 f(x) = f(x - h) - f(x), \quad (\text{A.10})$$

et

$$\Delta_h^2 f(x) = f(x - 2h) - 2f(x - h) + f(x). \quad (\text{A.11})$$

**Définition 10 (Modules de continuité)** Soit  $f \in L_2(\mathbb{R})$ . Les modules de continuités  $\omega^1$  et  $\omega^2$  sont définis pour tout  $x \geq 0$  par :

$$\omega_2^1(f, x) = \sup_{|h| < x} \|\Delta_h^1 f\|_2, \quad (\text{A.12})$$

et

$$\omega_2^2(f, x) = \sup_{|h| < x} \|\Delta_h^2 f\|_2. \quad (\text{A.13})$$

**Définition 11 (Espace de Besov  $B_2^{s,q}$ )** L'espace de Besov  $B_2^{s,q}(\mathbb{R})$ ,  $1 \leq q \leq +\infty$  est l'espace des fonctions  $f$  appartenant à  $W_2^n(\mathbb{R})$  et telles que :

$$\omega_2^2(f^{(n)}, x) = \epsilon(x)x^\alpha. \quad (\text{A.14})$$

où,

- $n$  est tel que  $s = n + \alpha$ , avec  $n \in \{0, 1, 2, \dots\}$  et  $0 < \alpha \leq 1$ ,
- $\epsilon$  est une fonction définie sur  $[0, +\infty[$  et telle que  $\|\epsilon\|_q^* < \infty$ , avec

$$\|\epsilon\|_q^* = \begin{cases} (\int_0^{+\infty} |\epsilon(x)|^q \frac{dx}{x})^{1/q} & \text{si } 1 \leq q < \infty, \\ \text{supess}(\epsilon) & \text{si } q = \infty, \end{cases}$$

supess étant le supremum essentiel, défini par

$$\text{supess}(\epsilon) = \inf\{c \in \mathbb{R}^+ : |\epsilon(x)| \leq c \text{ (pp)}\},$$

(pp) désignant presque partout.

La norme sur l'espace de Besov  $B_2^{s,q}$ , induite par la norme de l'espace de sobolev associé est définie par :

$$\|f\|_2^{s,q} = \|f\|_{W_2^n} + \left\| \frac{\omega_2^2(f^{(n)}, x)}{x^\alpha} \right\|_q^*. \quad (\text{A.15})$$

**Définition 12 (Opérateur sur un espace de Hilbert)** Soit  $\mathcal{H}$ , un espace de Hilbert. La donnée d'une application  $T : u \in \mathcal{H} \mapsto Tu \in \mathcal{H}$  définit un opérateur sur  $\mathcal{H}$ .

**Définition 13 (Opérateur borné)** Un opérateur est dit borné s'il transforme tout ensemble borné en un ensemble borné.

Un opérateur linéaire  $S$  est alors borné s'il est borné sur la boule unité. En effet, supposons qu'il existe une constante réelle  $K > 0$  tel que :

$$\|f\| \leq 1 \Rightarrow \|Sf\| \leq K.$$

Comme pour tout  $h \in \mathcal{H}^*$ , on a  $\|\frac{h}{\|h\|}\| = 1$ , on en déduit que  $\|S(\frac{h}{\|h\|})\| \leq K$ .  $S$  étant linéaire, on en conclut que

$$\|S(h)\| \leq K\|h\|. \quad (\text{A.16})$$

Cette dernière inégalité assure que l'image par  $S$  de toute partie bornée de  $\mathcal{H}$  est aussi une partie bornée. On peut alors définir la norme :

**Définition 14 (Norme d'opérateurs linéaires bornés)** On associe à l'espace des opérateurs linéaires bornés sur un espace de Hilbert, la norme  $||| \cdot |||$  définie par

$$|||S||| = \sup_{\|f\| \neq 0} \frac{\|Sf\|}{\|f\|}. \quad (\text{A.17})$$

**Proposition 7** *S étant un opérateur linéaire borné sur un espace de Hilbert  $\mathcal{H}$ , on a*

$$|||S||| = \sup_{\|f\|=1} \|Sf\| = \sup_{\|f\|\leq 1} \|Sf\|. \quad (\text{A.18})$$

**Preuve :** Soit  $f \in \mathcal{H}^*$ . Posons  $g = \frac{f}{\|f\|}$ . On a  $\|g\| = 1$ .

$$\frac{\|Sf\|}{\|f\|} = \frac{\|S(g\|f\|)\|}{\|f\|} = \|Sg\|.$$

D'autre part, soit  $f \in \mathcal{H}^*$  tel que  $\|f\| < 1$ . Pour  $g = \frac{f}{\|f\|}$ , on a :

$$\|Sf\| = \|f\| \cdot \|Sg\| \leq \|Sg\|.$$

■

Le théorème suivant caractérise l'inversion d'opérateurs linéaires bornés.

**Théorème 8 (Inversion d'opérateurs bornés)** *Soit S un opérateur linéaire et borné sur  $\mathcal{H}$ . S'il existe une constante  $A > 0$  telle que  $|||S||| > A$ , alors S est inversible et son inverse  $S^{-1}$  est tel que :*

$$|||S^{-1}||| \leq \frac{1}{A} \quad (\text{A.19})$$

**Preuve :** Voir [Daubechies \[1992\]](#), entre autres.

■

## A.2 Représentations dans les Espaces de Hilbert

Soit  $\Omega$  est un ensemble quelconque d'indices.

**Définition 15 (Frame)** *Une sequence  $\{\phi_n\}_{n \in \Omega}$  de vecteurs d'un espace de Hilbert  $\mathcal{H}$  est une frame de  $\mathcal{H}$  s'il existe deux constantes  $A > 0$  et  $B < \infty$  telles que*

$$\forall f \in \mathcal{H} \quad A\|f\|^2 \leq \sum_{n \in \Omega} |\langle f, \phi_n \rangle|^2 \leq B\|f\|^2. \quad (\text{A.20})$$

Considérons une frame  $\{\phi_n\}_{n \in \Omega}$  sur  $\mathcal{H}$ , et définissons un opérateur T sur  $\mathcal{H}$  par

$$T: f \in \mathcal{H} \longmapsto Tf = \sum_{n \in \Omega} \langle f, \phi_n \rangle \phi_n. \quad (\text{A.21})$$

T est un opérateur linéaire sur  $\mathcal{H}$ . On l'appelera opérateur de frames. On a

$$\langle Tf, f \rangle = \sum_{n \in \Omega} |\langle f, \phi_n \rangle|^2,$$

et donc :

$$\forall f \in \mathcal{H}, \quad A\|f\|^2 \leq \langle Tf, f \rangle \leq B\|f\|^2. \quad (\text{A.22})$$

La norme de l'opérateur de frames T est telle que  $A \leq |||T||| \leq B$  (voir annexe B pour la preuve). Il est donc borné et satisfait aux hypothèses du Théorème 8. L'opérateur T est donc inversible, et son inverse  $T^{-1}$  vérifie précisément,

$$\frac{1}{B} \leq |||T^{-1}||| \leq \frac{1}{A}. \quad (\text{A.23})$$

### A.2.1 Caractérisation d'une fonction par ses coefficients de frames

Soit  $f \in \mathcal{H}$ . On a :

$$f = T^{-1}Tf = T^{-1}\left(\sum_{n \in \Omega} \langle f, \phi_n \rangle \phi_n\right) = \sum_{n \in \Omega} \langle f, \phi_n \rangle T^{-1}\phi_n. \quad (\text{A.24})$$

Posons :

$$\widetilde{\phi}_n = T^{-1}\phi_n. \quad (\text{A.25})$$

La famille  $\{\widetilde{\phi}_n\}_{n \in \Omega}$  est aussi une frame appelée *frame duale* de  $\{\phi_n\}_{n \in \Omega}$ . L'expression (A.24) devient donc, compte tenu de l'Eq. (A.25) :

$$f = \sum_{n \in \Omega} \langle f, \phi_n \rangle \widetilde{\phi}_n. \quad (\text{A.26})$$

C'est le *développement par frames* de la fonction  $f \in \mathcal{H}$ , par *projection sur la frame duale*. Le développement par *projection sur la frame* s'écrit :

$$f = \sum_{n \in \Omega} \langle f, \widetilde{\phi}_n \rangle \phi_n. \quad (\text{A.27})$$

Dans ce développement, les coefficients  $(\langle f, \widetilde{\phi}_n \rangle)_n$  sont appelés *coefficients de frames*. La représentation par frames de  $f \in \mathcal{H}$  consiste donc à projeter  $f$  sur une frame, les coefficients de la représentation étant les produits scalaires avec les vecteurs de la frame duale. La donnée d'une frame sur un espace de Hilbert revient à munir cet espace d'un système générateur Hilbertien garantissant une représentation stable, puisque, d'après l'Eq. (A.20), la suite des coefficients de frames est bornée au sens de la norme  $\ell^2(\Omega)$  ( $\ell^2(\Omega)$  est l'espace des suites réelles ou complexes  $(x_n)_{n \in \Omega}$  vérifiant  $\sum_{n \in \Omega} |x_n|^2 < \infty$ ). Cette représentation peut être redondante car les vecteurs de frame ne sont pas forcément indépendants (au sens du produit scalaire considéré dans l'espace de Hilbert).

### A.2.2 Frame ajustée - Base de Riesz

**Définition 16 (Frame ajustée)** Une frame est dite *ajustée* ou *serrée* si les constantes  $A$  et  $B$  sont égales. On dit aussi parfois qu'elle est *tendue*.

Pour une frame ajustée, en faisant  $A = B$  dans l'Eq. (A.22), on a  $\langle Tf, f \rangle = A\|f\|^2$ , pour tout  $f \in \mathcal{H}$ ; ce qui est équivalent à  $\langle Tf - Af, f \rangle = 0$ . Si  $\text{Id}$  désigne l'opérateur identité, on a donc  $\langle (T - A.\text{Id})f, f \rangle = 0$ , ce qui est encore équivalent à  $\langle (T - A.\text{Id})f, g \rangle = 0$  pour tout  $f, g \in \mathcal{H}$ . D'où

$$\|T - A.\text{Id}\| = 0$$

L'opérateur  $T - A.\text{Id}$  est donc l'opérateur nul. On en déduit que

$$T = A.\text{Id}. \quad (\text{A.28})$$

Par conséquent, l'inverse  $T^{-1}$  de  $T$  est défini par

$$T^{-1} = \frac{1}{A}.\text{Id}, \quad (\text{A.29})$$

et on en conclut que :

$$\widetilde{\phi}_n = \frac{1}{A} \phi_n. \quad (\text{A.30})$$

Le développement par frame ajustée de  $f$  s'écrit donc

$$f = \frac{1}{A} \sum_{n \in \Omega} \langle f, \phi_n \rangle \phi_n. \quad (\text{A.31})$$

Une fonction  $f \in \mathcal{H}$  peut donc être représentée un développement du type donné par l'Eq. (A.31) sur une frame ajustée.

**Définition 17 (Base de Riesz)** Une base de Riesz est une frame dont les vecteurs sont linéairement indépendants.

Lorsque la frame est une base de Riesz, la frame duale est également une base de Riesz et on a  $\langle \phi_m, \widetilde{\phi}_n \rangle = \frac{1}{A} \delta(m-n)$ . Le système “frames - frame duale” est donc un système biorthogonal de vecteurs. Pour une base de Riesz normalisée, on a  $A = B = 1$ , et le développement de  $f$  sur cette base orthonormale Hilbertienne est alors :

$$f = \sum_{n \in \Omega} \langle f, \phi_n \rangle \phi_n. \quad (\text{A.32})$$

### A.2.3 Représentations en ondelettes

Les Transformations en Ondelettes sont des représentations Hilbertiennes qui possèdent la spécificité suivante : les vecteurs de la base de Hilbert sur laquelle se fait la projection sont générés à partir d'une seule fonction appartenant à  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  et appelée fonction d'ondelette.

**Définition 18 (Ondelette)** Une fonction  $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ , régulière et bien localisée, est appelé ondelette si :

$$\int_0^{+\infty} \frac{|\mathcal{F}\psi(\omega)|^2}{\omega} d\omega < +\infty. \quad (\text{A.33})$$

Par “bien localisée”, on veut que  $\psi$  soit une fonction à support compact ou à décroissance rapide et on impose en général la condition :  $t \mapsto t\psi(t) \in L^1(\mathbb{R})$ . Une ondelette est dite à  $m+1$  moments nuls si :

$$\int_{\mathbb{R}} t^k \psi(t) dt = 0 \quad \text{pour} \quad k = 0, 1, \dots, m. \quad (\text{A.34})$$

Les notions de Transformations en Ondelettes biorthogonales (respectivement orthogonales) se rapportent aux notions de bases de Riesz (respectivement Base de Riesz normalisée). Une Transformations en Ondelettes est redondante lorsque les fonctions de la frame d'ondelettes ne sont pas linéairement indépendants.



# ANNEXE B Bornes de l'opérateur de frames

**Proposition 8** L'opérateur de frames  $\mathbb{T} : f \in \mathcal{H} \mapsto \sum_{n \in \Omega} \langle f, \phi_n \rangle \phi_n$ , vérifie :

$$A \leq |||\mathbb{T}||| \leq B, \quad (\text{B.1})$$

où  $A$  et  $B$  sont les constantes définies par  $A\|f\|^2 \leq \sum_{n \in \Omega} |\langle f, \phi_n \rangle|^2 \leq B\|f\|^2$ , pour tout  $f \in \mathcal{H}$ .

**Rappel :** Si  $U : (f, g) \mapsto U(f, g)$  est une forme bilinéaire, la forme  $f \mapsto U(f, f)$  est appelée forme quadratique associée à  $U$ .

**Preuve :** [de la Proposition 8]. Soit la forme quadratique :

$$\mathcal{Q} : f \in \mathcal{H} \mapsto \langle \mathbb{T}f, f \rangle \in \mathbb{C}. \quad (\text{B.2})$$

$\mathcal{Q}$  est la forme positive (car  $\mathcal{Q}(f) \geq A\|f\|^2 \geq 0$ ) associée à la forme bilinéaire :

$$U : (f, g) \in \mathcal{H} \times \mathcal{H} \mapsto \langle \mathbb{T}f, g \rangle \in \mathbb{C}.$$

D'après l'Eq. (A.22), on a, pour tout  $f$  tel que  $\|f\| = 1$ ,

$$A \leq \langle \mathbb{T}f, f \rangle = \mathcal{Q}(f) \leq B.$$

On en déduit que :

$$A \leq |||\mathcal{Q}||| \leq B. \quad (\text{B.3})$$

La Proposition 8 est alors démontrée si  $|||\mathbb{T}||| = |||\mathcal{Q}|||$ . Ce qui est l'objet de la proposition suivante. ■

**Proposition 9** L'opérateur de frames  $\mathbb{T}$  est tel que :

$$|||\mathbb{T}||| = |||\mathcal{Q}|||, \quad (\text{B.4})$$

où  $\mathcal{Q}$  la forme quadratique positive donnée par l'Eq. (B.2).

**Preuve :** Elle procède en 2 étapes. Nous allons d'abord montrer que  $|||\mathbb{T}||| = |||U|||$ . Ensuite, nous montrerons que  $|||U||| = |||\mathcal{Q}|||$ .



1. Montrons que :

$$|||T||| = |||U|||. \quad (\text{B.5})$$

– On a :

$$|||U||| = \sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\langle Tf, g \rangle| \leq \sup_{\|f\| \leq 1} \|Tf\| = |||T|||.$$

Donc :

$$|||U||| \leq |||T|||. \quad (\text{B.6})$$

– D'autre part, on a :

$$|||U||| = \sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\langle Tf, g \rangle| \geq \sup_{\substack{\|f\| \leq 1 \\ \|g\|=1}} |\langle Tf, g \rangle|.$$

En particulier, pour  $g = \frac{Tf}{\|Tf\|}$ , on a  $\|g\| = 1$ , et donc :

$$|||U||| \geq \sup_{\|f\| \leq 1} |\langle Tf, \frac{Tf}{\|Tf\|} \rangle| = \sup_{\|f\| \leq 1} \|Tf\| = |||T|||.$$

Ainsi :

$$|||U||| \geq |||T|||. \quad (\text{B.7})$$

La relation (B.5) se déduit de (B.6) et de (B.7).

2. Montrons alors que :

$$|||U||| = |||\mathcal{Q}|||. \quad (\text{B.8})$$

– On a :

$$|||\mathcal{Q}||| = \sup_{\|f\| \leq 1} |\langle Tf, f \rangle| \leq \sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\langle Tf, g \rangle| = |||U|||.$$

Donc :

$$|||\mathcal{Q}||| \leq |||U|||. \quad (\text{B.9})$$

– On veut maintenant montrer que  $|||U||| \leq |||\mathcal{Q}|||$ . Soit  $\Re U(f, g)$ , la partie réelle du complexe  $U(f, g)$ . Montrons que :

$$\Re U(f, g) = \frac{1}{4} \{ \mathcal{Q}(f+g) - \mathcal{Q}(f-g) \}.$$

On a :

$$\begin{cases} \mathcal{Q}(f+g) = \langle T(f+g), f+g \rangle, \\ \mathcal{Q}(f-g) = \langle T(f-g), f-g \rangle. \end{cases}$$

Par décomposition de ces produits scalaires, on obtient :

$$\mathcal{Q}(f+g) - \mathcal{Q}(f-g) = 2\{\langle Tf, g \rangle + \langle Tg, f \rangle\} = 4\Re \langle Tf, g \rangle,$$

d'où :

$$\Re U(f, g) = \frac{1}{4} \{ \mathcal{Q}(f+g) - \mathcal{Q}(f-g) \}. \quad (\text{B.10})$$

On a ainsi :

$$|\Re U(f, g)| \leq \frac{1}{4} \{|\mathcal{Q}(f+g)| + |\mathcal{Q}(f-g)|\},$$

et donc,

$$|\Re U(f, g)| \leq \frac{1}{4} |||\mathcal{Q}||| \{ \|f+g\|^2 + \|f-g\|^2 \}.$$

Or, d'après l'identité du parallélogramme, on a :

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2),$$

et on obtient donc

$$|\Re U(f, g)| \leq \frac{1}{2} |||\mathcal{Q}||| \{ \|f\|^2 + \|g\|^2 \}. \quad (\text{B.11})$$

Considérons alors les nombres complexes  $U(f, g) = \langle \mathbb{T}f, g \rangle$ . Si à chaque couple  $(f, g)$ , avec  $\|f\| \leq 1$  et  $\|g\| \leq 1$  ; on peut trouver un couple  $(f^*, g^*)$ , avec  $\|f^*\| \leq 1$  et  $\|g^*\| \leq 1$  ; tel que le produit scalaire  $\langle \mathbb{T}f^*, g^* \rangle$  soit réel et de valeur absolue égale au module du nombre complexe  $\langle \mathbb{T}f, g \rangle$ , alors on aurait

$$\sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\langle \mathbb{T}f, g \rangle| = \sup_{\substack{\|f^*\| \leq 1 \\ \|g^*\| \leq 1}} |\langle \mathbb{T}f^*, g^* \rangle|,$$

et le sup serait donc aussi atteint sur les  $|\Re U(f, g)|$ . Prenons alors :

$$f^* = f \text{ et } g^* = g.e^{i\theta} \text{ avec } \theta = \text{Arg}(\langle \mathbb{T}f, g \rangle).$$

On a :

$$\|f^*\| = \|f\| \text{ et } \|g^*\| = \|g\|.$$

En outre, puisque

$$\langle \mathbb{T}f, g \rangle = |\langle \mathbb{T}f, g \rangle| e^{i\theta},$$

on a :

$$\langle \mathbb{T}f, g \rangle e^{-i\theta} = |\langle \mathbb{T}f, g \rangle|.$$

Or, par définition,

$$\langle \mathbb{T}f, g \rangle e^{-i\theta} = \langle \mathbb{T}f, g.e^{i\theta} \rangle = \langle \mathbb{T}f^*, g^* \rangle.$$

On en conclut donc que

$$\sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\langle \mathbb{T}f, g \rangle| = \sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} |\Re \langle \mathbb{T}f, g \rangle|,$$

et ainsi

$$|||U||| = \sup_{\|f\| \leq 1, \|g\| \leq 1} |\Re U(f, g)|. \quad (\text{B.12})$$

Compte tenu de l'Eq. (B.11), on obtient

$$|||U||| \leq \frac{1}{2} |||\mathcal{Q}||| \sup_{\substack{\|f\| \leq 1 \\ \|g\| \leq 1}} \{ \|f\|^2 + \|g\|^2 \},$$

et finalement

$$|||U||| \leq |||\mathcal{Q}|||. \quad (\text{B.13})$$

L'Eq. (B.8) se déduit alors de l'Eq. (B.9) et de l'Eq. (B.13). Finalement, à partir de Eqs (B.5) et (B.8), on a donc

$$|||\mathbb{T}||| = |||U||| = |||\mathcal{Q}|||. \quad (\text{B.14})$$

■

---

# Index

- Auto-Régressif (AR), 26–32, 38, 39  
Autocorrélation, 8, 9, 19–21, 91, 94  
Distribution gaussienne généralisée, 37–41, 71, 72, 85  
Fonction de seuillage, 2, 46, 50, 53–55, 57, 60–67, 69–71, 73, 74, 77, 83, 89  
Fonction SSBS, 47, 71, 73–88, 90–95  
Ondelettes, 1–3, 7–17, 19, 26, 29–31, 34–36, 41, 43, 46, 47, 49, 50, 57, 61, 62, 64, 66–72, 74, 82–95  
    *M*-TPOD, 7–17, 19–25, 33–36, 41, 42, 49, 50, 85, 93  
    2-TPOD, 14, 17, 26  
    3-TPOD, 14, 17, 18  
    TOD, 7–9, 11, 14, 49–53, 62, 64–66, 70, 78, 79, 81, 85–87  
    TPOD, 7–9, 11, 14, 22, 25–29, 37–41, 50, 86–88  
Parcimonie, 1–3, 7, 45, 46, 49, 50, 57–59, 61, 62, 64, 68–72, 76, 85, 86, 89, 91, 92, 94  
Stationnaire, 8–10, 19–24, 33–36, 41, 42, 46, 86, 91, 93  
SURE, 78–82, 90, 95

## Résumé

Les transformées en ondelettes et par paquets d'ondelettes se sont largement imposées dans l'analyse et la résolution de problèmes liés aux sciences et techniques de l'ingénierie. Cet essor est dû principalement à deux propriétés spécifiques qui résultent des décompositions sur les bases d'ondelettes : la parcimonie de représentation pour les signaux réguliers et réguliers par morceaux ; la tendance à transformer un processus aléatoire stationnaire en séquences de processus Gaussiens décorrélés. Ces deux propriétés sont observées expérimentalement et justifient de nombreux traitements effectués sur les coefficients obtenus par projection sur les bases d'ondelettes. Cependant, les résultats théoriques précisant la corrélation et la distribution asymptotiques des coefficients de paquets d'ondelettes sont de natures significativement différentes. La première partie de cette thèse propose une analyse de ces résultats théoriques, met en lumière la complexité du problème posé, et montre les résultats qu'on l'on peut obtenir à partir des critères de convergence. Les problèmes traités concernent à la fois les fonctions d'autocorrélation et les distributions asymptotiques des coefficients de paquets d'ondelettes de processus stationnaires. Les résultats annoncés conduisent à des schémas plus réalistes qui sont, de surcroît, soutenus par des résultats expérimentaux. La seconde partie de cette thèse analyse les propriétés de parcimonie des transformées à base d'ondelettes et propose l'utilisation de seuils associés à une mesure de parcimonie. Ces seuils établissent un lien entre la détection à l'estimation non-paramétrique, lien que nous résumons par la formule *bien détecter pour mieux estimer*. Ces seuils unifient les seuils minimax et universels, qui correspondent ainsi à des seuils de détection associés à des degrés différents de parcimonie. D'autre part, cette thèse unifie également les fonctions de seuillage de base (seuillage dur et doux) de l'estimation non-paramétrique par l'introduction d'une famille de fonctions à atténuation de type sigmoïde : les fonctions SSBS. Les fonctions de seuillage dur et doux sont alors des fonctions SSBS dégénérées, les fonctions SSBS non-dégénérées étant des fonctions régulières et à degré d'atténuation flexible. Les performances de l'estimation non-paramétrique utilisant les seuils de détection et les fonctions SSBS sont également analysées, tant du point de vue de l'erreur quadratique moyenne que de celui de la qualité visuelle en débruitage d'images.

## Mots clés

Ondelettes, Paquets d'Ondelettes, Processus aléatoires, Théorèmes Limites, Estimation Non-Paramétrique.

## Abstract

Wavelets and wavelet packet transforms are widely used for analyzing and solving problems related to science and engineering techniques. This growth is mainly due to two specific properties that result from decompositions on wavelet bases : the sparse representation of regular and piecewise regular signals, the tendency to transform a stationary random process into quasi decorrelated and Gaussian sequences. Both properties are observed experimentally and justify many treatments carried out on the coefficients obtained by projection on wavelet bases. However, the theoretical results describing the asymptotic autocorrelations and distributions associated with wavelet packet coefficients are not always of the same nature. The first part of the thesis provides these results. The results obtained thus lead to more realistic schemes for the asymptotic autocorrelation functions and distributions of wavelet packet coefficients and are supported by numerous experimental results. The second part of the thesis analyzes sparsity properties of wavelet bases and proposes the use of sparse-descriptive based thresholds. These thresholds establish a link between non-parametric detection and estimation in the sense that good detection leads to better estimation. The detection thresholds obtained also unify the minimax and universal thresholds in the sense that these thresholds correspond to detection thresholds associated with different sparsity degrees. On the other hand, this thesis also unifies basic soft and hard thresholding functions by introducing a new family of smooth sigmoid based shrinkage functions : the SSBS functions. In fact, the hard and soft thresholding functions are degenerated SSBS functions, in contrast to non-degenerated SSBS functions which are smooth with very flexible attenuation degrees. In addition, the performance of the non-parametric estimation using detection thresholds and SSBS functions is analyzed for natural image denoising, *via* the mean square error and visual quality assessment.

## Keywords

Wavelets, Wavelet packets, Stochastic processes, Limit theorems, Non-Parametric Estimation.